



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 10687

TO: Celine Qian
Location: CM1/11C01&11E12
Art Unit: 1636
Friday, October 24, 2003

Case Serial Number: 09/830703

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 24, 2003, 07:25:57 ; Search time 7756 Seconds
(without alignments)
17168.751 Million cell updates/sec

Title: US-09-830-703-1
Perfect score: 3255
Sequence: 1 ctcacgaggggaaggggga.....aaaaaaaaaaaaaaaa 3255

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2885711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pi.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.or.*
21: em.ov.*
22: em.pat.*
23: em.ph.*
24: em.pl.*
25: em.ro.*
26: em.sts.*
27: em.un.*
28: em.vi.*
29: em.vt.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3255	100.0	3255	6	AX089963 Sequence
2	3255	100.0	3255	6	AX136828 Sequence
3	3255	100.0	3255	6	AX089880 Sequence
4	3253.4	100.0	3255	6	AX089881 Sequence
5	3253.4	100.0	3255	6	AX089882 Sequence
6	3253.4	100.0	3255	6	AX136845 Sequence
7	3253.4	100.0	3255	6	AX136846 Sequence
8	3253.4	100.0	3255	6	AX136847 Sequence
9	3243	99.6	3254	6	AX089877 Sequence
10	3243	99.6	3254	6	AX136842 Sequence
11	3241	99.6	3253	6	AX089876 Sequence
12	3241	99.6	3253	6	AX089878 Sequence
13	3241	99.6	3253	6	AX136841 Sequence
14	3241	99.6	3253	6	AX136843 Sequence
15	3226	99.1	3226	10	AF250293 Mus muscu
16	3226	99.1	3226	6	AX089873 Sequence
17	3075	94.5	3170	6	AX136838 Sequence
18	3007	92.4	3136	6	AX089872 Sequence
19	3007	92.4	3136	6	AX136837 Sequence
20	2956	90.8	3092	6	AX089879 Sequence
21	2956	90.8	3092	6	AX136844 Sequence
22	2821	86.7	3043	6	AX089875 Sequence
23	2821	86.7	3043	6	AX136840 Sequence
24	2763	84.9	3014	6	AX089869 Sequence
25	2763	84.9	3014	6	AX136834 Sequence
26	2602.8	80.0	2895	6	AX089870 Sequence
27	2602.8	80.0	2895	6	AX136835 Sequence
28	2571	79.0	2918	6	AX089874 Sequence
29	2571	79.0	2918	6	AX136839 Sequence
30	2260.2	69.4	2558	6	AX089871 Sequence
31	2260.2	69.4	2558	6	AX136836 Sequence
32	1811.6	55.7	32621	2	AC116854 Mus muscu
33	1644	50.5	1644	10	AB019558 Mus muscu
34	1373.6	42.2	1564	10	AB019878 Rattus no
35	1352	41.5	1576	10	AF215434 Rattus no
36	1236.2	37.9	1456	10	AF168004 Rattus no
37	1226.2	37.9	1398	10	AF252234 Rattus no
38	1144.8	35.2	1470	10	AF381277 Rattus no
39	1130.4	34.4	1578	10	AF381278 Rattus no
40	1072	32.9	1554	10	AF381280 Rattus no
41	1058.4	32.5	1374	10	AF281281 Rattus no
42	1054	32.4	2960	6	AX584224 Sequence
43	1054	32.4	2960	9	AB009973 Homo sapi
44	1052.4	32.3	2560	6	AX584226 Sequence
45	1004	30.8	24887	2	AC135826 Rattus no

ALIGNMENTS

RESULT 1
LOCUS: AX089863
DEFINITION: Sequence 1 from Patent WO016176.
ACCESSION: AX089863
VERSION: AX089863.1 GI:13443960
KEYWORDS: Mus musculus (house mouse)
SOURCE: Mus musculus
ORGANISM: Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS: Luebbert, H.
TITLE: Transgenic animal model for neurodegenerative diseases
JOURNAL: Patent: WO 01/6176-A ; 03-VAR-2001;

FEATURES		Biofrontera Pharmaceuticals GmbH (DE)	
source	Location/Qualifiers		
	1..3255		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
BASE COUNT	852 a 809 g 776 t		
ORIGIN			
	Query Match 100.0%; Score 3255; DB 6; Length 3255;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 3255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 CTCAGCGAGGGAAGGGGAGGAGGCTGGATGACCTAAACCTTGACAGAAACGCTGGTGGG	60	
DB	1 CTCAGCGAGGGAAGGGGAGGAGGCTGGATGACCTAAACCTTGACAGAAACGCTGGTGGG	60	
QY	61 AGGCTCGGCGGGCGCCAGTCCCGGCTAGTCTCTCGACCCGAGCCACGACCGGCC	120	
DB	61 AGGCTCGGCGGGCGCCAGTCCCGGCTAGTCTCTCGACCCGAGCCACGACCGGCC	120	
QY	121 CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGGCTTCCCACTGGAGGTCG	180	
DB	121 CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGGCTTCCCACTGGAGGTCG	180	
QY	181 ACTCTGACACCCAGCATCTTGCAGCTCAAGGAAGTGGTCTTAAGCGACAGAGGGTCCAG	240	
DB	181 ATTCTGACACCCAGCATCTTGCAGCTCAAGGAATGGTGTCTAAGCGACAGAGGGTCCAG	240	
QY	241 CTGACCACTGCGTGTGATTTTCCGGGAAGAGCTTCGGAATCACTGATCGTTCAAA	300	
DB	241 CTGACCACTGCGTGTGATTTTCCGGGAAGAGCTTCGGAATCACTGATCGTTCAAA	300	
QY	301 ACTGTGACCTGGAACACAGAGTATTGTACACATAGTACGAGACCAAGGAGAGTGC	360	
DB	301 ACTGTGACCTGGAACACAGAGTATTGTACACATAGTACGAGACCAAGGAGAGTGC	360	
QY	361 ATGAACAAATGCACTTGGAGGGGAGCAACCCAGAGCACTTCAGAGGCTCCATATGG	420	
DB	361 ATGAACAAATGCACTTGGAGGGGAGCAACCCAGAGCACTTCAGAGGCTCCATATGG	420	
QY	421 AGTCCAGAGCTTGACACAGTGCACCTGAGCAGCCATACCCCTGCGGTGACTCTGG	480	
DB	421 AGTCCAGAGCTTGACACAGTGCACCTGAGCAGCCATACCCCTGCGGTGACTCTGG	480	
QY	481 GGCTGGCGGTCTTCTGACACAGAGTAAAGAGGATTCAGAAAGCAGCAGAGTCCAG	540	
DB	481 GGCTGGCGGTCTTCTGACACAGAGTAAAGAGGATTCAGAAAGCAGCAGAGTCCAG	540	
QY	541 TTAACCCACCTACACAGCTTTTTCATCTACTGCAAGAGGCCCTGCCACAGGTCACG	600	
DB	541 TTAACCCACCTACACAGCTTTTTCATCTACTGCAAGAGGCCCTGCCACAGGTCACG	600	
QY	601 CTGGAAGCTCCGAGTTCAGTTCGGACCTTGCACCAAGCAACCTCAGCTTGGCCAGG	660	
DB	601 CTGGAAGCTCCGAGTTCAGTTCGGACCTTGCACCAAGCAACCTCAGCTTGGCCAGG	660	
QY	661 GCCATCTTCTGGACGATGCTTAAATCCAAACCGGATGAGTGTGAGTGCAGTCTC	720	
DB	661 GCCATCTTCTGGACGATGCTTAAATCCAAACCGGATGAGTGTGAGTGCAGTCTC	720	
QY	721 CAGACTGCCCTGGACAGAGCTCAATTTTCTTAAATGTGGAGCAGACCCAACTCAG	780	
DB	721 CAGACTGCCCTGGACAGAGCTCAATTTTCTTAAATGTGGAGCAGACCCAACTCAG	780	
QY	781 ACAAGGACACGTGGTATGCTTTGAACCTGATCACCAGCCAGCGCAGCATCCCTTSCA	840	
DB	781 ACAAGGACACGTGGTATGCTTTGAACCTGATCACCAGCCAGCGCAGCATCCCTTSCA	840	
QY	841 TAGCGTGCACAGATGTCAGAGCCCTGCTGCTGCTTCTCAGTGTAAACACCTGACGTA	900	
DB	841 TAGCGTGCACAGATGTCAGAGCCCTGCTGCTGCTTCTCAGTGTAAACACCTGACGTA	900	

QY	901 TCTGTTTGACATGTTTCCACTGTTACTGTGTGTCACAGACTCAACGATCGGCAGTTGTGTC	960
DB	901 TCTGTTTGACATGTTTCCACTTGTATTGTGTGTCACAGACTCAACGATCGGCAGTTGTGTC	960
QY	961 ACGATGCTCAACTTGGCTACTCCCTGCGGTGTGTAGCTGTCCCACTCCCTGATTA	1020
DB	961 ACGATGCTCAACTTGGCTACTCCCTGCGGTGTGTAGCTGTCCCACTCCCTGATTA	1020
QY	1021 AAGASCTCACTACATTCAGGATCCTTGGAGAGAGCACTACCTAGGTACACAGCATATG	1080
DB	1021 AAGAGCTCCATCACTTTCAGGATCCTTGGAGAGAGCACTAGTACACTAGTACACAGCATATG	1080
QY	1081 GGGCCGAGGAATGGCTGCTGCAAAATGGAGAGTGTGCTGTGCCCCCGCTCTCTGGTGGAG	1140
DB	1081 GGGCCGAGGAATGGCTGCTGCAAAATGGAGAGTGTGCTGTGCCCCCGCTCTCTGGTGGAG	1140
QY	1141 CTGGACTGCTACTGACAGCGGCCAGAGAGAAAGTCACTCTCGAAGAGGGGCAACGGCCTG	1200
DB	1141 CTGGACTGCTACTGACAGCGGCCAGAGAGAAAGTCACTCTCGAAGAGGGGCAACGGCCTG	1200
QY	1201 GCTGGGGTTTTTTTCTGCGGAGCTGTAAAGGAAGCATACCATGAAGGAGGATTTGGCACT	1260
DB	1201 GCTGGGGTTTTTTTCTGCGGAGCTGTAAAGGAAGCATACCATGAAGGAGGATTTGGCACT	1260
QY	1261 CACTGCTCGAACCTTCAGAGAGCACTTCTCAGGCTTACAGGTTGGACAAAGAGAGCGCTG	1320
DB	1261 CACTGCTCGAACCTTCAGAGAGCACTTCTCAGGCTTACAGGTTGGACAAAGAGAGCGCTG	1320
QY	1321 AGCAAGCTCGTGGAGAGAGCCCTCAAGGAAGACATCAAGAGAGAGCCCAAGCCCTTGTCT	1380
DB	1321 AGCAAGCTCGTGGAGAGAGCCCTCAAGGAAGACATCAAGAGAGAGCCCAAGCCCTTGTCT	1380
QY	1381 CTCGCTGCAAGTCCCAATTCGAAAAGAGAGAGATGTATGCACTAGAGTCTCTCAGC	1440
DB	1381 CTCGCTGCAAGTCCCAATTCGAAAAGAGAGAGATGTATGCACTAGAGTCTCTCAGC	1440
QY	1441 CCCAGTGGAACTGGAGTGTGGAACTGTGGCTGTGAGTGGAAACCGAGCCCTGATG	1500
DB	1441 CCCAGTGGAACTGGAGTGTGGAACTGTGGCTGTGAGTGGAAACCGAGCCCTGATG	1500
QY	1501 GAGATCACTGTTTGTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGACGACAACTCAA	1560
DB	1501 GAGATCACTGTTTGTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGACGACAACTCAA	1560
QY	1561 GGGAACTCCGAGAGTCTCTACCTTCTTACGCTTCT	

Db	306:	ATTGTGTGAGTGTGAATAGTGGGATAAAAGTGGATGTCAGAGAGTGGAAATCAAAACCTC	312
QY	3121	TGCAAAAGCAATCTTCTCTCTCTGTGGAAGTGTATTAGAAATACCTGAAAGTCTGTGTGTG	3180
Db	3121	TGCAAAAGCAATCTTCTCTCTCTGTGGAAGTGTATTAGAAATACCTGAAAGTCTGTGTGTG	3180
QY	3181	TGTTGTATCCCGAGCTGTCAATCAATAAAGACCGAGCTGTGCAATGAAMAAAAAAMAAA	3240
Db	3181	TGTTGTATCCCGAGCTGTCAATCAATAAAGACCGAGCTGTGCAATGAAMAAAAAAMAAA	3240
QY	3241	AAAAAAMAAAAAAA 3255	
Db	3241	AAAAAAMAAAAAAA 3255	
RESULT 2			
AX136828			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3255;			
Conservative			
0;			
Mismatches			
0;			
Indels			
0;			
Gaps			
0;			
1			
CTCAGCAGCGGGAAGGGGAGGAGGCGCTGGATGACTAAACCTTGACAGAAACCGCTGGTGGG			60
1			
CTCAGCAGCGGGAAGGGGAGGAGGCGCTGGATGACTAAACCTTGACAGAAACCGCTGGTGGG			60
61			
AGGCTCGGCGGGCGCCAGTCGCCCGGTAGTCCCTCTCGACCCCGACGACACCGCC			120
61			
AGGCTCGGCGGGCGCCAGTCGCCCGGTAGTCCCTCTCGACCCCGACGACACCGCC			120
121			
CGGTGACCATGATAGTGTGTTGTGAGCTTCAACTCCAGCTATGCTTCCGAGTGGAGTCTG			180
121			
CGGTGACCATGATAGTGTGTTGTGAGCTTCAACTCCAGCTATGCTTCCGAGTGGAGTCTG			180
181			
ATTCGACACCGAGCATCTTGCAGCTTCAGGAATGTTGTTCTAAGCGACAGAGGGGTTCCAG			240
181			
ATTCGACACCGAGCATCTTGCAGCTTCAGGAATGTTGTTCTAAGCGACAGAGGGGTTCCAG			240
241			
CTGACCAAGCTGGTGTGATTCTTGCGGGAGAGGCTTCCGAATCACTGACGGTTCCAAA			300
241			
CTGACCAAGCTGGTGTGATTCTTGCGGGAGAGGCTTCCGAATCACTGACGGTTCCAAA			300
301			
ACTGTGACTCGAAACACAGAGATATTGTACATATGTCAGAGACCCAGAGAGAGATC			360
301			
ACTGTGACTCGAAACACAGAGATATTGTACATATGTCAGAGACCCAGAGAGAGATC			360
361			
ATGAACAAATGCACTCTGAGGGGAGCAACCCAGGACCTCAGAGGGCTCATATGG			420
361			
ATGAACAAATGCACTCTGAGGGGAGCAACCCAGGACCTCAGAGGGCTCATATGG			420
421			
AGTCCAGGAGCTTGACACAGTGGACCTTGAGCAGCATACCTCGCGGTGACTCTGTGG			480

Dbb 421 AGTCAGGAGCTTGACAGAGTGGACCTGAGCAGCCATACCTGCGGTGGACTCTCTGG 480
QY 481 GCTGGCGGTCAATCTGGACACAGACAGTAGAGGGATTACAGAGCAGCCAGAGSTCCAG 540
Db 481 GCTGGCGGTCAATCTGGACACAGACAGTAGAGGGATTACAGAGCAGCCAGAGSTCCAG 540
QY 541 TTAAACCCACCTACAAACAGCTTTTTCATCTACTGCAAAAGGCCCTGGCCACAAAGTCCAGC 600
Db 541 TTAAACCCACCTACAAACAGCTTTTTCATCTACTGCAAAAGGCCCTGGCCACAAAGTCCAGC 600
QY 601 CTGGAAGCTCCGAGTTCAGTGTGGCACTGTGCAAAACAAAGCAACCTCACCTTGGCCAGG 660
Db 601 CTGGAAGCTCCGAGTTCAGTGTGGCACTGTGCAAAACAAAGCAACCTCACCTTGGCCAGG 660
QY 661 GCCCATCTTGTGGAGCAGATGCTTAATTCGAAACCGGATGAGTGGTGGAGTCC 720
Db 661 GCCCATCTTGTGGAGCAGATGCTTAATTCGAAACCGGATGAGTGGTGGAGTCC 720
QY 721 CAGACTGCCCTGGAACCCAGAGCTGAATTTTCTTAAATGTGGAGCAGCCCAACCTCAG 780
Db 721 CAGACTGCCCTGGAACCCAGAGCTGAATTTTCTTAAATGTGGAGCAGCCCAACCTCAG 780
QY 781 ACNAGGACAGTGGGTAGTGTGACCTGATCACCAGCAAGCGGAGCATCCCTTGCA 840
Db 781 ACNAGGACAGTGGGTAGTGTGACCTGATCACCAGCAAGCGGAGCATCCCTTGCA 840
QY 841 TAGCGTGACAGATGTGAGAGCCCTGCTCTTCCAGTGTAAACCCAGCTCACGTGA 900
Db 841 TAGCGTGACAGATGTGAGAGCCCTGCTCTTCCAGTGTAAACCCAGCTCACGTGA 900
QY 901 TCTGTTGGACTGTTCCACTTGTATTTGTGTACAAAGACTCAACGATCGGCAGTTGTCC 960
Db 901 TCTGTTGGACTGTTCCACTTGTATTTGTGTACAAAGACTCAACGATCGGCAGTTGTCC 960
QY 961 AGCATGCTCAACTGGCTACTCCCTGCGTGTAGCTGGGTGCTCCCAACTCCCTGATTA 1020
Db 961 AGCATGCTCAACTGGCTACTCCCTGCGTGTAGCTGGGTGCTCCCAACTCCCTGATTA 1020
QY 1021 AAGAGTCCATCACTTCCAGTTCAGTGTGAGAGAGCAGTACACTAGTACCCAGATG 1080
Db 1021 AAGAGTCCATCACTTCCAGTTCAGTGTGAGAGAGCAGTACACTAGTACCCAGATG 1080
QY 1081 GGGCCGAGAAATGCGTGTGCAAAATGGGAGGTGTGCTGTGCCCCCGTCTGGCTGTGGAG 1140
Db 1081 GGGCCGAGAAATGCGTGTGCAAAATGGGAGGTGTGCTGTGCCCCCGTCTGGCTGTGGAG 1140
QY 1141 CTGGACTGCTACTGCAACAGGCGCCAGAGGAAAGTCACTGCGAAGGGGCAACGCGCTGG 1200
Db 1141 CTGGACTGCTACTGCAACAGGCGCCAGAGGAAAGTCACTGCGAAGGGGCAACGCGCTGG 1200
QY 1201 GCTGGGGTTTGTCTGCGGGAGCTGTAGGAGGATACATGAAGGGATTTGGACT 1260
Db 1201 GCTGGGGTTTGTCTGCGGGAGCTGTAGGAGGATACATGAAGGGATTTGGACT 1260
QY 1261 CACTGCTCGAACCCCTCAGGAGCACTTCTCAGGCGCTACAGGTTGACAAAGAGCCGCTG 1320
Db 1261 CACTGCTCGAACCCCTCAGGAGCACTTCTCAGGCGCTACAGGTTGACAAAGAGCCGCTG 1320
QY 1321 AGCAAGCTCGTGGAGGAGGCTCCAGGAAACCATCAAGAGACCCAGGCTTGTG 1380
Db 1321 AGCAAGCTCGTGGAGGAGGCTCCAGGAAACCATCAAGAGACCCAGGCTTGTG 1380
QY 1381 CTGCTGTCAACGTGCCAATTTGAATAACGGAGGATGTATGCATATGAATGTCTCAGC 1440
Db 1381 CTGCTGTCAACGTGCCAATTTGAATAACGGAGGATGTATGCATATGAATGTCTCAGC 1440
QY 1441 CCAGTGCAGGTGGAGTGGTCTCGAATGTGGCTGTGAGTGAACCGAGCCCTGCATGG 1500
Db 1441 CCAGTGCAGGTGGAGTGGTCTCGAATGTGGCTGTGAGTGAACCGAGCCCTGCATGG 1500
QY 1501 GAGATCACTGGTTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGAGCGACACCTCAA 1560
Db 1501 GAGATCACTGGTTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGAGCGACACCTCAA 1560

QY 1561 GGGAACTCCGAGATTCCTACCTTCCTTACCGATTCCTTCTCTCATATTAAGCA 1620
Db 1561 GGGAACTCCGAGATTCCTACCTTCCTTACCGATTCCTTCTCTCATATTAAGCA 1620
QY 1621 CATTAATCGGACACACAAAACAGAGGTGCAGATACAGAGCAGCCCTAGATCCTTTC 1680
Db 1621 CATTAATCGGACACACAAAACAGAGGTGCAGATACAGAGCAGCCCTAGATCCTTTC 1680
QY 1681 TAGGCAACCCACAGAAACACAGACACCCGCTGCCCCAGGGGAGGAGCAGCTTTCAGC 1740
Db 1681 TAGGCAACCCACAGAAACACAGACACCCGCTGCCCCAGGGGAGGAGCAGCTTTCAGC 1740
QY 1741 CTCGGCTCACTCGAATGTGAGAGCTTAGATGAGGTGCACCTTTGGTTGGATCTGTA 1800
Db 1741 CTCGGCTCACTCGAATGTGAGAGCTTAGATGAGGTGCACCTTTGGTTGGATCTGTA 1800
QY 1801 GAAGCAAGTGGAGTGGGAGTGTTCAGAGTGTTCGCCAGCCCTGGGTAAAGTAA 1860
Db 1801 GAAGCAAGTGGAGTGGGAGTGTTCAGAGTGTTCGCCAGCCCTGGGTAAAGTAA 1860
QY 1861 CACCTCTGAGATTCAGAGACACCTTGAGATCTCAGGAGCCCTGCTCATGTAGTA 1920
Db 1861 CACCTCTGAGATTCAGAGACACCTTGAGATCTCAGGAGCCCTGCTCATGTAGTA 1920
QY 1921 ATCATCTATTCCCAAGGGGCCCTGCGAGTAGTCAAAACTATTGTTCATCCGCCCAAT 1980
Db 1921 ATCATCTATTCCCAAGGGGCCCTGCGAGTAGTCAAAACTATTGTTCATCCGCCCAAT 1980
QY 1981 CCTATCTTACAAATGGCTGATGAGATACACCCCTCTCTGTACTATCAGCTTATC 2040
Db 1981 CCTATCTTACAAATGGCTGATGAGATACACCCCTCTCTGTACTATCAGCTTATC 2040
QY 2041 AACCAAGTGGAGACTAGGAAAGCTAATGATGGCAGACTGCTTAAATCGCAGGAGGA 2100
Db 2041 AACCAAGTGGAGACTAGGAAAGCTAATGATGGCAGACTGCTTAAATCGCAGGAGGA 2100
QY 2101 CTGAGAGCCAAACCTTCTGCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 2160
Db 2101 CTGAGAGCCAAACCTTCTGCTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 2160
QY 2161 TTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 TTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY 2221 GCTGGAAGTTTAGCTTCACTGTGAATTTAAACAGAGAAAGTGCCTATAAAGGGGCGTTT 2280
Db 2221 GCTGGAAGTTTAGCTTCACTGTGAATTTAAACAGAGAAAGTGCCTATAAAGGGGCGTTT 2280
QY 2281 TTAAGAGCAATCCCATGATCTGCGCAATGCTTAACAGAGGCTCAGAAACACATGT 2340
Db 2281 TTAAGAGCAATCCCATGATCTGCGCAATGCTTAACAGAGGCTCAGAAACACATGT 2340
QY 2341 TTAAGAGGAGCAGCTTCCGACCATCTGAATGAGATGATGCTGACCCCTTCCACCACA 2400
Db 2341 TTAAGAGGAGCAGCTTCCGACCATCTGAATGAGATGATGCTGACCCCTTCCACCACA 2400
QY 2401 AGTGGGAGCAGCTTCCGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 2401 AGTGGGAGCAGCTTCCGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
QY 2461 CCACCAGTGGTCTTACAGAGAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC 2520
Db 2461 CCACCAGTGGTCTTACAGAGAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC 2520
QY 2521 ATGATCACTCAAGTGTAACTTTCAAGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2580
Db 2521 ATGATCACTCAAGTGTAACTTTCAAGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2580
QY 2581 TGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2640
Db 2581 TGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2640

RESULT 4					
AX089881					
LOCUS					
DEFINITION	Sequence 19 from Patent WO01-6176.	3255 bp	DNA	linear	PAT 21-MAR-2003
ACCESSION	AX089881				
VERSION	AX089881.1	G:13443975			
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Luebbert H;				
AUTHORS	Transgenic animal model for neurodegenerative diseases				
TITLE	Patent: WO 016176-A 19 08-MAR-2003;				
JOURNAL	BioFrontiers Pharmaceuticals GmbH (DE)				
FEATURES	Location/Qualifiers				
source	1..3255				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10093"				
BASE COUNT	853 A 817 C 909 G 776 T				
ORIGIN					
	Query Match:	100.0%	Score 3253.4;	DB 6;	Length 3255;
	Best Local Similarity	100.0%;	Fred. No. 0;		
	Matches 3254;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1 CTCACGAGGGGAAGGGGAGAGCGCTTGGATGACTTAACCTTGACAGAAACCGCTGTGGG	60			
Db	1 CTCACGAGGGGAAGGGGAGAGCGCTTGGATGACTTAACCTTGACAGAAACCGCTGTGGG	60			
QY	61 AGGCTCGCGCGGGGCCAGTCGCCGCCTAGTGCCTCTCTGACC CGCAGCCACCCGCC	120			
Db	61 AGGCTCGCGCGGGGCCAGTCGCCGCCTAGTGCCTCTCTGACC CGCAGCCACCCGCC	120			
QY	121 CGGTGACCATGATAGTTGTTTCAGGTTCAACTCCAGCTATGSGCTTCCCAGTGGAGSGTCG	180			
Db	121 CGGTGACCATGATAGTTGTTTCAGGTTCAACTCCAGCTATGSGCTTCCCAGTGGAGSGTCG	180			
QY	181 ATTCTGACACACAGCATCTTGCAAGTCTCAAGNAGTGTGCTTAGCGACACAGGGGGTTCAG	240			
Db	181 ATTCTGACACACAGCATCTTGCAAGTCTCAAGNAGTGTGCTTAGCGACACAGGGGGTTCAG	240			
QY	241 CTGACCAGCTCGCTGTGATTTTTTCCGCGGAGGAGGTTCGGAATACCTTGACGSGTTCAAA	300			
Db	241 CTGACCAGCTCGCTGTGATTTTTTCCGCGGAGGAGGTTCGGAATACCTTGACGSGTTCAAA	300			
QY	301 ACTGTGACCTTGGACACACAGAGTATTGTACACATAGTACAGACACACCGAGGAGAAGTC	360			
Db	301 ACTGTGACCTTGGACACACAGAGTATTGTACACATAGTACAGACACACCGAGGAGAAGTC	360			
QY	361 ATGAACAATAATGTCATCTCGAGGGGACGAAACCCACAGACGCTCAGAGGGCTCCATATGGG	420			
Db	361 ATGAACAATAATGTCATCTCGAGGGGACGAAACCCACAGACGCTCAGAGGGCTCCATATGGG	420			
QY	421 AGTCACAGAGCTTGNACGAGTGGACCTGTGAGCGCCATACCCCTGCCGTGGAGCTCTGTGG	480			
Db	421 AGTCACAGAGCTTGNACGAGTGGACCTGTGAGCGCCATACCCCTGCCGTGGAGCTCTGTGG	480			
QY	481 GSGTCGGCGTCACTCTGACACACAGACAGTAGTAAGAGGGATTTCAGAAAGCAGCCAGAGTCCAG	540			
Db	481 GSGTCGGCGTCACTCTGACACACAGACAGTAGTAAGAGGGATTTCAGAAAGCAGCCAGAGTCCAG	540			
QY	541 TTAAACCCACCTTACAAACGCTTTTTCATCTACTGCAAGGCCCCCTGCCACAAAGGTCACGC	600			
Db	541 TTAAACCCACCTTACAAACGCTTTTTCATCTACTGCAAGGCCCCCTGCCACAAAGGTCACGC	600			
QY	601 CTGGAAGAGCTCCGAGTTTCAGTGTGSCACCTTGCAAAACGAAACCTCTGAGCTTGGGCCACG	660			
Db	601 CTGGAAGAGCTCCGAGTTTCAGTGTGSCACCTTGCAAAACGAAACCTCTGAGCTTGGGCCACG	660			

```
QY 1741 CTCTGGCTCACTCGAATGTGAGAGCTTAGAGTGGGTGACCTTTGGTTTGGATCTGTGA 1800
DB 1741 CTC-TGG-CTCACTCGAATGTGAGAGCTTAGAGTGGGTGACCTTTGGTTTGGATCTGTGA 1800
QY 1801 GAAGCCATGAGTGGAGTGGAGTGGTTTTCAGAGGTGTTTGGCCAGCCCTGGGTGAAGTAA 1860
DB 1801 GAAGCCATGAGTGGAGTGGAGTGGTTTTCAGAGGTGTTTGGCCAGCCCTGGGTGAAGTAA 1860
QY 1861 CACCTCTGAGGATCTCAGAAACACATCTGAGATCTCAGGAACGCTCTCTCATGTAGTA 1920
DB 1861 CACCTCTGAGGATCTCAGAAACACATCTGAGATCTCAGGAACGCTCTCTCATGTAGTA 1920
QY 1921 ATCATCTATTTCCAAAGGGCCCTCGCAGTAGTCAAAACTATTGTTTATCCCCCAAAAT 1980
DB 1921 ATCATCTATTTCCAAAGGGCCCTCGCAGTAGTCAAAACTATTGTTTATCCCCCAAAAT 1980
QY 1981 CCTATCTTTCAAAATGCTGATGAGATTACACCCCTCTGTGTACTAATGAGCTTATC 2040
DB 1981 CCTATCTTTCAAAATGCTGATGAGATTACACCCCTCTGTGTACTAATGAGCTTATC 2040
QY 2041 AACCAAGTGAAGCTAGGAAGCTAATTGGATGGCAGACTGCTTAAATCGCAGGGAGGA 2100
DB 2041 AACCAAGTGAAGCTAGGAAGCTAATTGGATGGCAGACTGCTTAAATCGCAGGGAGGA 2100
QY 2101 CTCAGAGCCAAACCTACTTCCGTTGCTTTCATTTATCTGCAACTTTGGAGTATCTCTACTGCA 2160
DB 2101 CTCAGAGCCAAACCTACTTCCGTTGCTTTCATTTATCTGCAACTTTGGAGTATCTCTACTGCA 2160
QY 2161 TTTTTCCTCCCTGAAAGATTAACAAAGTCTGCAATTTGGAGTATCTCTACTGCA 2220
DB 2161 TTTTTCCTCCCTGAAAGATTAACAAAGTCTGCAATTTGGAGTATCTCTACTGCA 2220
QY 2221 GCCTGGAAGTTAGCTTACCTGCAATTTAACAGAGAAAGTGCCTATAAAGGGGGGCTT 2280
DB 2221 GCCTGGAAGTTAGCTTACCTGCAATTTAACAGAGAAAGTGCCTATAAAGGGGGGCTT 2280
QY 2281 TTAAGAGACATCCCATGATGCTGGGCAATGCTTAACAGAGAGTCAAGAAACACATGT 2340
DB 2281 TTAAGAGACATCCCATGATGCTGGGCAATGCTTAACAGAGAGTCAAGAAACACATGT 2340
QY 2341 TTAAGAGAGACATCCCATGATGCTGGGCAATGCTTAACAGAGAGTCAAGAAACACATGT 2400
DB 2341 TTAAGAGAGACATCCCATGATGCTGGGCAATGCTTAACAGAGAGTCAAGAAACACATGT 2400
QY 2401 AGTGGGGACATCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
DB 2401 AGTGGGGACATCTGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
QY 2461 CCACCCAGTGTCTCAGACAGGGCAATACACACACAGAGTACGCTTCAGATCAAC 2520
DB 2461 CCACCCAGTGTCTCAGACAGGGCAATACACACACAGAGTACGCTTCAGATCAAC 2520
QY 2521 ATGATCAGACTCAAGTGTAAATCTTTCAAGGTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2580
DB 2521 ATGATCAGACTCAAGTGTAAATCTTTCAAGGTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2580
QY 2581 TGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 2640
DB 2581 TGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 2640
QY 2641 GCCTAGAGCTAAAATCATATAGAAATGATCTTATCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 2700
DB 2641 GCCTAGAGCTAAAATCATATAGAAATGATCTTATCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 2700
QY 2701 GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
DB 2701 GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
QY 2761 GAAAGAGCCACTCTGCAAGCAGCAGTCAAGGAGTCTGATGTCACCCACTATTTTTTTT 2820
DB 2761 GAAAGAGCCACTCTGCAAGCAGCAGTCAAGGAGTCTGATGTCACCCACTATTTTTTTT 2820
QY 2821 CCATCATTTGTCTTGCCTCTGCCTCTCTTCCACACCCCTGTGAGTAAATCGCATTTGGGAAG 2880
```

```
DB 2821 CCATCATTTGTCTTGCCTCTGCCTCTCTTCCACACCCCTGTGAGTAAATCGCATTTGGGAAG 2880
QY 2861 CCAGSACAAATGTTTGTCTGTTCTGTTCTTGGGTAAAGGGACTCCCTGAAAGCTCTGTGGTCTCT 2940
DB 2861 CCAGSACAAATGTTTGTCTGTTCTGTTCTTGGGTAAAGGGACTCCCTGAAAGCTCTGTGGTCTCT 2940
QY 2941 CCAGTATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3000
DB 2941 CCAGTATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3000
QY 3001 ATTCTTAAATPAAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCAACAGCTGC 3060
DB 3001 ATTCTTAAATPAAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCAACAGCTGC 3060
QY 3061 ATTCTGTGAGTGTGAATAGTGGGATAAAGTGGTGTGAGAGAGTGGAAATCAACAGCTC 3120
DB 3061 ATTCTGTGAGTGTGAATAGTGGGATAAAGTGGTGTGAGAGAGTGGAAATCAACAGCTC 3120
QY 3121 TGCAAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
DB 3121 TGCAAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
QY 3181 TGGTGTGAGTGTGAATAGTGGGATAAAGTGGTGTGAGAGAGTGGAAATCAACAGCTC 3240
DB 3181 TGGTGTGAGTGTGAATAGTGGGATAAAGTGGTGTGAGAGAGTGGAAATCAACAGCTC 3240
QY 3241 AAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3300
DB 3241 AAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3300
LOCUS AX089882 3255 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 20 from Patent WO9816176.
ACCESSION AX089882
VERSION AX089882.1 GI:3443976
KEYWORDS Mus musculus (house mouse);
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Luebbert H.
AUTHORS Transgenic animal model for neurodegenerative diseases
TITLE Patent: WO 9816176-A 20 08-MAR-2001.
JOURNAL Biofrontiers Pharmaceuticals GmbH (DE)
FEATURES Location/Qualifiers
source 1..3255
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="axon:1009c"
BASE COUNT 853 a 818 c 808 g 776 t
ORIGIN
Query Match 100.0%; Score 3253.4; DR 6; Length 3255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCAGCAGGCGCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 1 CTCAGCAGGCGCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
QY 61 AGGCTCGGGCGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 AGGCTCGGGCGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CGGTGACCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 121 CGGTGACCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 181 ATCTGTGACACAGCATCTTGCAGCTCAAGGAAGTGGTGTCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
```


QY	2401	AGTGGGACACCTCTGCAATATCTGCTCCCTCCTCTGCTGTTTAAGCCCCAGGGAGCCCA	2460
DB	2401	AGTGGGACACCTCTGCAATATCTGCTCCCTCCTCTGCTGTTTAAGCCCCAGGGAGCCCA	2460
QY	2461	CCACCCAGTGGTCTACAGACAGGGCAATACACACACCAACAATAGCCTTCAGATCAAC	2520
DB	2461	CCACCCAGTGGTCTACAGACAGGGCAATACACACACCAACAATAGCCTTCAGATCAAC	2520
QY	2521	ATGCATCACACTCAAGTGTAACTTTTCAAGGTTTTCTTTTCCTTTTCCTTGTTTTTAT	2580
DB	2521	ATGCATCACACTCAAGTGTAACTTTTCAAGGTTTTCTTTTCCTTTTCCTTGTTTTTAT	2580
QY	2581	TGTTTTGCTTTTGCTTTTTTTTTTTTTTTTTTTTTTTTGGTGGTGGTGGGGCTACCAACTTGAG	2640
DB	2581	TGTTTTGCTTTTGCTTTTTTTTTTTTTTTTTTTTTTTTGGTGGTGGTGGGGCTACCAACTTGAG	2640
QY	2641	GCCTAGAGCCAAAATCATATAGAAATGATGTTATCTTTGGTGGTGGAGGAAGCCAGCT	2700
DB	2641	GCCTAGAGCTAAAAATCATATAGAAATGATGTTATCTTTGGTGGTGGAGGAAGCCAGCT	2700
QY	2701	GGCCTAAGTTTACACTTTTGTCCCACTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT	2760
DB	2701	GGCCTAAGTTTACACTTTTGTCCCACTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT	2760
QY	2761	GAAAGACCACTGTCAAGCAGCAGTCAAGAGTCTGATGTCAACCCCTTGGGAG	2820
DB	2761	GAAAGACCACTGTCAAGCAGCAGTCAAGAGTCTGATGTCAACCCCTTGGGAG	2820
QY	2821	CCATCATTTGCTTGCCTCTGCCTCCCTCCACACCCCTGTGAGCTAATCGATGGGAG	2880
DB	2821	CCATCATTTGCTTGCCTCTGCCTCCCTCCACACCCCTGTGAGCTAATCGATGGGAG	2880
QY	2881	CCAGGACAAATGTTTGTCTTTCTGCTTTGGGTAAAGGAGCTCCCTGAGCTCTGTGCTCT	2940
DB	2881	CCAGGACAAATGTTTGTCTTTCTGCTTTGGGTAAAGGAGCTCCCTGAGCTCTGTGCTCT	2940
QY	2941	CCAGTATGGTCCCTTTTCTCTCTTAAACAGATGCAATATGTTTCTTCAAGATACAAATAGTG	3000
DB	2941	CCAGTATGGTCCCTTTTCTCTCTTAAACAGATGCAATATGTTTCTTCAAGATACAAATAGTG	3000
QY	3061	ATTCTTTAAATTAACCCAAAAGACAGCATCCACAGTGTGTGAGCATGAATCAAGCTCTC	3120
DB	3061	ATTCTTTAAATTAACCCAAAAGACAGCATCCACAGTGTGTGAGCATGAATCAAGCTCTC	3120
QY	3121	ATTGTGTGAGTGAATAGTGGGATAAAAAGTGGATGTGAGAGAGTGGAAATCAAACTTC	3180
DB	3121	ATTGTGTGAGTGAATAGTGGGATAAAAAGTGGATGTGAGAGAGTGGAAATCAAACTTC	3180
QY	3181	TGCAAGCAATCTTTCTCTTTCTGGAAGTGTATTAGAAATACCTGAACTCTGTGTGTG	3240
DB	3181	TGCAAGCAATCTTTCTCTTTCTGGAAGTGTATTAGAAATACCTGAACTCTGTGTGTG	3240
QY	3241	TGCTGGTACCCAGACTGTCATCAATAAAGACCCAGACTGTCMAATGAAAAAATAAAAAA	3300
DB	3241	TGCTGGTACCCAGACTGTCATCAATAAAGACCCAGACTGTCMAATGAAAAAATAAAAAA	3300
QY	3361	AAAAAATAAAAAA 3255	
DB	3361	AAAAAATAAAAAA 3255	

RESULT 6	
AXI36845	
LOCUS	AXI36845
DEFINITION	Sequence 18 from Patent EP081225.
ACCESSION	AXI36845
VERSION	AXI36845.1 Gz:14273210
KEYWORDS	
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

Db 3001 ATTCTTAAATAAACCCAAAGACAGGCATCCACAGTGTGTGAGCATGAATCACAGCCTGC 3063

QY 3061 ATTGTGTAGTGTGANTAGTGGGATTAAGTGGATGTGAGAGAGTGGAAATCAAACTTC 3120

Db 3061 ATTGTGTAGTGTGANTAGTGGGATTAAGTGGATGTGAGAGAGTGGAAATCAAACTTC 3120

QY 3121 TGCAGAGCAATCTTCTCTTCTTCTGTAAGTGTATTAAAGAAATACCTGAAGTCTGTGTGTG 3190

Db 3121 TGCAGAGCAATCTTCTCTTCTTCTGTAAGTGTATTAAAGAAATACCTGAAGTCTGTGTGTG 3190

QY 3181 TGGTGTATCCAGAGCTGTCAATCAATAAGACCCAGAGCTGTCAATGAATAAAAAA 3240

Db 3181 TGGTGTATCCAGAGCTGTCAATCAATAAGACCCAGAGCTGTCAATGAATAAAAAA 3240

QY 3241 AAAAAAAAAAAAAA 3255

Db 3241 AAAAAAAAAAAAAA 3255

RESULT 7

AX136846

LOCUS 3255 bp DNA linear PAT 30-VAY-2001

DESCRIPTION Sequence 19 from Patent EPI08-225.

ACCESSION AX136846

VERSION AX136846.1 GI:14273211

KEYWORDS

SOURCE Mus musculus (house mouse);

ORGANISM Mus musculus

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. Luebert, H.

AUTHORS Transgenic animal model for neurodegenerative diseases

TITLE Patent: EP 1081225-A 19 07-MAR-2001;

JOURNAL Biofrontiers Pharmaceuticals AG (DE)

FEATURES

source Location/Qualifiers

1..3255

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

BASE COUNT 853 a 817 c 809 g 776 t

ORIGIN

Query Match 100.0%; Score 3253.4; DB 6; Length 3255;

Best Local Similarity 100.0%; Pred. No. 5;

Matches 3254; Conservative 3; Mismatches 1; Indels 6; Gaps 0;

QY 1 CTCAGCGAGGGGAAGGGGAGGAGGCTGGATGACTAAACCTTGACAGAAACGCTGGTGGG 60

Db 1 CTCAGCGAGGGGAAGGGGAGGAGGCTGGATGACTAAACCTTGACAGAAACGCTGGTGGG 60

QY 61 AGCTCGGGGGGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120

Db 61 AGCTCGGGGGGGGGCCAGTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120

QY 121 CGGTGACCATGATAGT 180

Db 121 CGGTGACCATGATAGT 180

QY 181 ATTCTGACACCATCTTTCAGCTCAAGGAAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240

Db 181 ATTCTGACACCATCTTTCAGCTCAAGGAAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240

QY 241 CTGACCATGCTGT 300

Db 241 CTGACCATGCTGT 300

QY 301 ACTGTGACCTGGAAACACAGATATTGTACATAGTACAGACCCAGGAGGAGTGC 360

Db 301 ACTGTGACCTGGAAACACAGATATTGTACATAGTACAGACCCAGGAGGAGTGC 360

QY 361 ATGAACAAATGATCTGGAGGGAGACCCAGACCTCAGAGGCTTCATATGGG 420

Db 361 ATGAACAAATGATCTGGAGGGAGACCCAGACCTCAGAGGCTTCATATGGG 420

QY 421 AGTCAGGAGCTGTACACAGTGGACCTGAGACGACCATACCTGCTCCGTGGACTCTCTGG 480

Db 421 AGTCAGGAGCTGTACACAGTGGACCTGAGACGACCATACCTGCTCCGTGGACTCTCTGG 480

QY 481 GCTCGCGGTGATCTGAGACACAGACGATGAGAGGATTCAGAGAGACCGAGAGTCCAG 540

Db 481 GCTCGCGGTGATCTGAGACACAGACGATGAGAGGATTCAGAGAGACCGAGAGTCCAG 540

QY 541 TTAACCCACCTCACACAGCTTTTTCATCTACTGCAAAAGGCGCTTGCACAAAGTCCAGC 600

Db 541 TTAACCCACCTCACACAGCTTTTTCATCTACTGCAAAAGGCGCTTGCACAAAGTCCAGC 600

QY 601 CTGGAAGCTCCGAGTTCAGTGTGCACTGCAACACAGCAACCTCAGCTTGGCCAGG 660

Db 601 CTGGAAGCTCCGAGTTCAGTGTGCACTGCAACACAGCAACCTCAGCTTGGCCAGG 660

QY 661 GCTCATCTTGTGGAGCATGTCTTAATTCACACCGSATGAGTGTGAGTGGCAGTCTC 720

Db 661 GCTCATCTTGTGGAGCATGTCTTAATTCACACCGSATGAGTGTGAGTGGCAGTCTC 720

QY 721 CAGACTGCTCGGAACACAGACTGAATTTTCTTTAAATGTGAGAGACACCCACCTCAG 780

Db 721 CAGACTGCTCGGAACACAGACTGAATTTTCTTTAAATGTGAGAGACACCCACCTCAG 780

QY 781 ACAAGGACACGTCGGTAGCTTTGAACCTGTACACACACAGCGCGAGCATCCCTGCA 840

Db 781 ACAAGGACACGTCGGTAGCTTTGAACCTGTACACACACAGCGCGAGCATCCCTGCA 840

QY 841 TAGCTGTACAGATGTGAGAGACCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

Db 841 TAGCTGTACAGATGTGAGAGACCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

QY 901 TCTGTTTGGAGCTTTCCACTTGTATTGTGTGACAGACTCAACGATCGAGGTTTCTC 960

Db 901 TCTGTTTGGAGCTTTCCACTTGTATTGTGTGACAGACTCAACGATCGAGGTTTCTC 960

QY 961 AGCATGCTCAACTTGGTACTTCCCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020

Db 961 AGCATGCTCAACTTGGTACTTCCCTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020

QY 1021 AAGACTGCTCACTTCCAGATCTTGGAGAGAGCAGTACATAGTATCAGCAGTATG 1080

Db 1021 AAGACTGCTCACTTCCAGATCTTGGAGAGAGCAGTACATAGTATCAGCAGTATG 1080

QY 1081 GGGCGAGGAAATGGCTGTGCAAAATGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140

Db 1081 GGGCGAGGAAATGGCTGTGCAAAATGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140

QY 1141 CTGGACTGCTACCTGACAGCGGCTGAGAGAGATACCTGCGAGAGGAGGAGCGGCTGG 1200

Db 1141 CTGGACTGCTACCTGACAGCGGCTGAGAGAGATACCTGCGAGAGGAGGAGCGGCTGG 1200

QY 1201 GCTCGGGT 1260

Db 1201 GCTCGGGT 1260

QY 1261 CACTGCTGACACCTCAGGAGCCTTCTCAGGCTTACAGGCTGAGCAAAAGAGCGCTG 1320

Db 1261 CACTGCTGACACCTCAGGAGCCTTCTCAGGCTTACAGGCTGAGCAAAAGAGCGCTG 1320

QY 1321 AGCAAGCTCGGTGGAGGAGGCTTCAAGGAGAACCATCAAGAGACCCACAGCCTTCTC 1380

Db 1321 AGCAAGCTCGGTGGAGGAGGCTTCAAGGAGAACCATCAAGAGACCCACAGCCTTCTC 1380

QY 1381 CTCGCTGCAACGTCGCAATTCGAAAGAACGAGAGATGTAAGCATGAAAGTCTCTCAGC 1440

Db 1381 CTCGCTGCAACGTCGCAATTCGAAAGAACGAGAGATGTAAGCATGAAAGTCTCTCAGC 1440

QY 1441 CCGACTGCAACGTCGAGT 1500

Db 1441 CCGACTGCAACGTCGAGT 1500

```
QY 1501 GAGATCACTGGTTGAGCTGTAGAGAGAGATGTCACTTGGCCCTGGAGCGACAACCTCAA 1560
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 GAGATCACTGGTTGAGCTGTAGAGAGAGATGTCACTTGGCCCTGGAGCGACAACCTCAA 1560
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 GGGAACTCCGAGATTCCTACCTTCTTCCAGCCATTCCTTCTCGATGCAATAAAGCA 1620
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 GGGAACTCCGAGATTCCTACCTTCTTCCAGCCATTCCTTCTCGATGCAATAAAGCA 1620
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1621 CATAAATGCCGACACACAAACACAGGCTGCAGATTTACAGAGAGAGCCCTTAGATCCTTTC 1680
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1621 CATAAATGCCGACACACAAACACAGGCTGCAGATTTACAGAGAGAGCCCTTAGATCCTTTC 1680
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1681 TAGGCAACCCACAGAAACACACAGCACCCGCTGGCCCGAGGGGAGAGCACTTTCAGC 1740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1681 TAGGCAACCCACAGAAACACACAGCACCCGCTGGCCCGAGGGGAGAGCACTTTCAGC 1740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1741 CTCGGGCTCACTGGAATGTCAGAGCTTAGATGAGGCTGCACCTTTTGGTTGATTCGTA 1800
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1741 CTCGGGCTCACTGGAATGTCAGAGCTTAGATGAGGCTGCACCTTTTGGTTGATTCGTA 1800
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 SAAGCCATGAGTGAGGTGGGAAGTGTCTTCCAGGGTGTTCGCCACGCCCTGGGTAAAGTAA 1960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 SAAGCCATGAGTGAGGTGGGAAGTGTCTTCCAGGGTGTTCGCCACGCCCTGGGTAAAGTAA 1960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 CACCTGTAGGATTTCTCAGAAAGCACACTTGAGATCTGAGGAAGCGTCTCATGTAGTA 1920
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 CACCTGTAGGATTTCTCAGAAAGCACACTTGAGATCTGAGGAAGCGTCTCATGTAGTA 1920
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1921 ATCATCTATTCCAAAGGGCCCTGCAGTAGTCAAAACCTATTGTATTCCGCCCAAAAT 1980
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1921 ATCATCTATTCCAAAGGGCCCTGCAGTAGTCAAAACCTATTGTATTCCGCCCAAAAT 1980
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1981 CCTATCTTTACAAATGCTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC 2040
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1981 CCTATCTTTACAAATGCTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC 2040
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2041 AACCAAGTGAGACCTAGGAAGCTAATGGATGGCAGACTGCTTAAATCGCAGGAGGA 2100
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2101 CTGAGAGCCAAACCTACTTCCGTTCCGTTTCATTATCTGCAACCTTTAGAAAGAAATGATC 2160
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2101 CTGAGAGCCAAACCTACTTCCGTTCCGTTTCATTATCTGCAACCTTTAGAAAGAAATGATC 2160
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2161 TTTTTCCTCCCTGAAAGATAACAAAGTCTGCAATTTGGTTGGAGTATTCCTACTGCA 2220
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2221 GCCTGGAGTGTAGCTTCACTGTGMAATTTAACAGAGAAAGTGCCTATAAGGGGGGGTTT 2280
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2221 GCCTGGAGTGTAGCTTCACTGTGMAATTTAACAGAGAAAGTGCCTATAAGGGGGGGTTT 2280
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2281 TTAAGACACAATCCCATGCTGCGCAATGCTAACCAACAGGCTCAAGAAACACAATGT 2340
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2281 TTAAGACACAATCCCATGCTGCGCAATGCTAACCAACAGGCTCAAGAAACACAATGT 2340
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2341 TTAGAGAGAGAGATCCCTCGACCATCTGAATGAGAGTATSCCTGACCCCTTCCACCACA 2400
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2341 TTAGAGAGAGAGATCCCTCGACCATCTGAATGAGAGTATSCCTGACCCCTTCCACCACA 2400
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2401 AGTGGGACACCTCTGATATCTGCTCCCTCTCTGTTAGCCGCCAGGAGCCCAT 2460
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2401 AGTGGGACACCTCTGATATCTGCTCCCTCTCTGTTAGCCGCCAGGAGCCCAT 2460
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2461 CCACCCAGTGTCTACAGACAGGGCAATACACACACCAAGATAGCCCTTCAGATCAAC 2520
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2461 CCACCCAGTGTCTACAGACAGGGCAATACACACACCAAGATAGCCCTTCAGATCAAC 2520
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2521 ATGCATCACACTCAAGTGTAACTTTTCAAGGTTTCTTCTCTTTTCTGTTTATTT 2580
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2521 ATGCATCACACTCAAGTGTAACTTTTCAAGGTTTCTTCTCTTTTCTGTTTATTT 2580
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 2581 TGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGG 2640
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2581 TGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGG 2640
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2641 GCCTAGAGCTTAAATATCATATAGAAATGATGTTATGTTGGTGTGAGGAAAGCCAGCT 2700
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2641 GCCTAGAGCTTAAATATCATATAGAAATGATGTTATGTTGGTGTGAGGAAAGCCAGCT 2700
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2701 GGCCTAAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCAGCTCCCAAAAT 2760
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2701 GGCCTAAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCAGCTCCCAAAAT 2760
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2761 SAAMAGACCACTGTCAAGGACAGCTCAGAGTCTGATGTCATCCATCCTATTTTTTTT 2820
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2761 SAAMAGACCACTGTCAAGGACAGCTCAGAGTCTGATGTCATCCATCCTATTTTTTTT 2820
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2821 CCATCATGTGCTTGGCTCTGCTCTCCCTCCACACCGCTGTGACGATTAATCGCATTTGGGAG 2880
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2821 CCATCATGTGCTTGGCTCTGCTCTCCCTCCACACCGCTGTGACGATTAATCGCATTTGGGAG 2880
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2881 CCAGSACAAATGTTTSCCTGTTCTGTTTGGTTAAAGGGACTCCCTGAAAGCTTGTGGCTCT 2940
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2881 CCAGSACAAATGTTTSCCTGTTCTGTTTGGTTAAAGGGACTCCCTGAAAGCTTGTGGCTCT 2940
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2941 CCAGTATGGTGGCTTTCCTTCTTACAGATGCTATGTTTCTTTCAGATACATATAGT 3000
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2941 CCAGTATGGTGGCTTTCCTTCTTACAGATGCTATGTTTCTTTCAGATACATATAGT 3000
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3001 ATTCTTAAATTAACCCAAAGACAGGCTCCACAGTGTGTGAGGATGAATCCACAGCTGC 3060
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3001 ATTCTTAAATTAACCCAAAGACAGGCTCCACAGTGTGTGAGGATGAATCCACAGCTGC 3060
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3061 ATTCTGTGAGTGTGAATAGTGGGATAAAGTGGATGTGACAGAGAGTGGAAATCAAACTC 3120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3061 ATTCTGTGAGTGTGAATAGTGGGATAAAGTGGATGTGACAGAGAGTGGAAATCAAACTC 3120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3121 TGCATAGCAATCTTCTCTTCTGTGTGAGTGTATTAAGAAATACCTTAAGTCTGTGTG 3180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3121 TGCATAGCAATCTTCTCTTCTGTGTGAGTGTATTAAGAAATACCTTAAGTCTGTGTG 3180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3181 TGGTGGTACCAGAGTGTCAATCAATTAAGACCCAGACTGTCAATGAAAAAANAANA 3240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3181 TGGTGGTACCAGAGTGTCAATCAATTAAGACCCAGACTGTCAATGAAAAAANAANA 3240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3241 AAAAAAANAANAANA 3255
DB 3241 AAAAAAANAANAANA 3255
```

```
RESULT 8
AX136847 AX136847 3255 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 20 from Patent EP1081225.
DEFINITION
ACCESSION AX136847
VERSION AX136847.1 G1:14273212
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
REFERENCE
1 Luebbert, H.
AUTHORS Luebbert, H.
TITLES Transgenic animal model for neurodegenerative diseases
JOURNAL Patent: EP 1081225-A 20 MAY-2001;
Biofrontiers Pharmaceuticals AG (DE)
FEATURES
source
1 3255
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10093"
BASE COUNT 853 a 818 c 908 g 776 t
ORIGIN
```

Query Match				: 100.0% ; Score 3253.4 ; DB 6 ; Length 3255 ;			
Best Local Similarity 100.0% ; Pred. No. 0 ;							
Matches 3254 ; Conservative 0 ; Mismatches 1 ; Indels 0 ; Gaps 0 ;							
QY	1	CTCAGCGAGGGAGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG	60				
DB	1	CTCAGCGAGGGAGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG	60				
QY	61	AGGCTCGGGGGGGCCAGTGGCCCGGTAGTTCCTTCGACCCGACGCCACCCGCC	120				
DB	61	AGGCTCGGGGGGGCCAGTGGCCCGGTAGTTCCTTCGACCCGACGCCACCCGCC	120				
QY	121	CGGTGACCATGATAGTGTGTGTGACAGCTTCAACTCCAGCTATGGCTTCCAGTGGAGTGG	180				
DB	121	CGGTGACCATGATAGTGTGTGTGACAGCTTCAACTCCAGCTATGGCTTCCAGTGGAGTGG	180				
QY	181	ATTCTGACACACGACATCTTTCGAGCTCAAGGAAAGTGGTTCGAGGAGGAGTTCGAG	240				
DB	181	ATTCTGACACACGACATCTTTCGAGCTCAAGGAAAGTGGTTCGAGGAGGAGTTCGAG	240				
QY	241	CTGACCAAGCTGCTGTGTATTTTTCGCGGAAGAGCTTCCGAATCACCTGACGTTCAAA	300				
DB	241	CTGACCAAGCTGCTGTGTATTTTTCGCGGAAGAGCTTCCGAATCACCTGACGTTCAAA	300				
QY	301	ACTGTGACCTGGAAACACAGAGTATTTTACATAGTACAGAGACACGAGGAGATC	360				
DB	301	ACTGTGACCTGGAAACACAGAGTATTTTACATAGTACAGAGACACGAGGAGATC	360				
QY	361	ATGAAACAAATCGATCTGGAGGGGACGAACCCGAGAGCACTTCAGAGGCTCCATATGG	420				
DB	361	ATGAAACAAATCGATCTGGAGGGGACGAACCCGAGAGCACTTCAGAGGCTCCATATGG	420				
QY	421	AGTCCAGGAGCTTGACACAGAGTGGACCTGAGGAGCCATACCTCGCGGTGGACTCTGG	480				
DB	421	AGTCCAGGAGCTTGACACAGAGTGGACCTGAGGAGCCATACCTCGCGGTGGACTCTGG	480				
QY	481	GGTGTGGGTGATCTGGACACAGAGTAAAGAGGATTCAGAGAGCCAGAGTCCAG	540				
DB	481	GGTGTGGGTGATCTGGACACAGAGTAAAGAGGATTCAGAGAGCCAGAGTCCAG	540				
QY	541	TTAAACCCACCTACACAGCTTTTTCATCTACTGCAAGGCGCCCTGGCACAGGTCCAGC	600				
DB	541	TTAAACCCACCTACACAGCTTTTTCATCTACTGCAAGGCGCCCTGGCACAGGTCCAGC	600				
QY	601	CTGAAAGCTCCGAGTTCAGTGTGGACCTTGAACAAAGCAACCTCACCTTGGCCACAG	660				
DB	601	CTGAAAGCTCCGAGTTCAGTGTGGACCTTGAACAAAGCAACCTCACCTTGGCCACAG	660				
QY	661	GCCCATCTTCTGGGACGATGCTTAATTCBAAACCGGATGATGGTGGTGGTGGTGGT	720				
DB	661	GCCCATCTTCTGGGACGATGCTTAATTCBAAACCGGATGATGGTGGTGGTGGTGGT	720				
QY	721	CAGACTGCTGCTGGAACAGAGTGAATTTTCTTTAAATGTGGAGCACACCACTCAG	780				
DB	721	CAGACTGCTGCTGGAACAGAGTGAATTTTCTTTAAATGTGGAGCACACCACTCAG	780				
QY	781	ACAGGACAGCTCGGTGATGCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840				
DB	781	ACAGGACAGCTCGGTGATGCTGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840				
QY	841	TAGCGTGACAGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	900				
DB	841	TAGCGTGACAGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	900				
QY	901	TCTGTTTGGACTGTTTCCACTGATGTTGTCACAGACTCAACAGTCCGAGTGGTCC	960				
DB	901	TCTGTTTGGACTGTTTCCACTGATGTTGTCACAGACTCAACAGTCCGAGTGGTCC	960				
QY	961	ACGATGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020				
DB	961	ACGATGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020				
QY	1021	AAGAGCTCCATCACTTCAGGATCCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG	1080				

DB	1021	AAGAGCTCCATCACTTCAGGATCCTTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG	1080				
QY	1081	GGGCTCGAGAAATGCGTGTCTGCAAAATGGGAGGTGTCTGTGTGCCCCCGTCTGCTGTG	1140				
DB	1081	GGGCTCGAGAAATGCGTGTCTGCAAAATGGGAGGTGTCTGTGTGCCCCCGTCTGCTGTG	1140				
QY	1141	CTGGACTGCTACTGAAACAGGSCCAGAGGAAAGTCACTGCGAGAGGGGCGACGCTGG	1200				
DB	1141	CTGGACTGCTACTGAAACAGGSCCAGAGGAAAGTCACTGCGAGAGGGGCGACGCTGG	1200				
QY	1201	GCTGCGGGTGTGTTTCTGCGGGAGCTGTAAAGGAAGCATACCATGAAAGGGATTGCGACT	1260				
DB	1201	GCTGCGGGTGTGTTTCTGCGGGAGCTGTAAAGGAAGCATACCATGAAAGGGATTGCGACT	1260				
QY	1261	CAGTGTGCAACCTTCAGSAGGACCTTCTCAGGCTCAGAGGCTGCAAGAGAGGCGGCTG	1320				
DB	1261	CAGTGTGCAACCTTCAGSAGGACCTTCTCAGGCTCAGAGGCTGCAAGAGAGGCGGCTG	1320				
QY	1321	AGCAAGCTCGCTGGGAGGAGGCTTCAAGGAAACCATCAAGAGAGCCACCAAGCCTTCTC	1380				
DB	1321	AGCAAGCTCGCTGGGAGGAGGCTTCAAGGAAACCATCAAGAGAGCCACCAAGCCTTCTC	1380				
QY	1381	CTGCTGCAAGCTGCGCAATTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1440				
DB	1381	CTGCTGCAAGCTGCGCAATTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1440				
QY	1441	CCAGTGTGAGGCTGGAGTGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1500				
DB	1441	CCAGTGTGAGGCTGGAGTGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1500				
QY	1501	GAGTCACTGGTGTGACGTGTAGAGAGAGATGTCTGCTGAGGAGGAGGAGGAGGAGG	1560				
DB	1501	GAGTCACTGGTGTGACGTGTAGAGAGAGATGTCTGCTGAGGAGGAGGAGGAGGAGG	1560				
QY	1561	GGGAAACTCGGAGATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1620				
DB	1561	GGGAAACTCGGAGATTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1620				
QY	1621	CATAAATGGGACACACAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1680				
DB	1621	CATAAATGGGACACACAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1680				
QY	1681	TAGGACACACAGAGAAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740				
DB	1681	TAGGACACACAGAGAAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1740				
QY	1741	CTGCTGCTCACTGGAATGTGAGAGGTTAGATGAGGAGGAGGAGGAGGAGGAGGAGG	1800				
DB	1741	CTGCTGCTCACTGGAATGTGAGAGGTTAGATGAGGAGGAGGAGGAGGAGGAGGAGG	1800				
QY	1801	GAAGCATGAGTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1860				
DB	1801	GAAGCATGAGTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1860				
QY	1861	CAGCTGTGAGGATTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1920				
DB	1861	CAGCTGTGAGGATTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1920				
QY	1921	ATCATCTATTTCCAAAGGCGCCCTGCGAGTGTGCAAAACTATTTTATCCCCCAAAAT	1980				
DB	1921	ATCATCTATTTCCAAAGGCGCCCTGCGAGTGTGCAAAACTATTTTATCCCCCAAAAT	1980				
QY	1981	CCTATCTTTACAAATGGTGTGATGAGATTAACAGGAGGAGGAGGAGGAGGAGGAGG	2040				
DB	1981	CCTATCTTTACAAATGGTGTGATGAGATTAACAGGAGGAGGAGGAGGAGGAGGAGG	2040				
QY	2041	AACCAAGTGAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2100				
DB	2041	AACCAAGTGAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2100				
QY	2101	AGGAGTGAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2160				
DB	2101	AGGAGTGAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2160				

2101 CTGAGAGCCAAACCTACTTCCGTTCCGTTTCATATTCTGCAACTTTAGAAAGAAATGATC 2160
2161 TTTTTCCTCCCTCAAAAGATTAACAAAGTCTGCAATTTGGTTTGAGTATTCTCTACTGCA 2220
2162 TTTTTCCTCCCTGAAAGATTAACAAAGTCTGCAATTTGGTTTGAGTATTCTCTACTGCA 2220
2221 GCCTGAGTTTACCTTCTACTGTGAATTTAAAGAGAAAGTCCCTATAAAGGGGGGGTTT 2290
2221 GCCTGAGTTTACCTTCTACTGTGAATTTAAAGAGAAAGTCCCTATAAAGGGGGGGTTT 2290
2281 TTAAGAGACAATCCCATGATGCTGGCCCAATGCTAACACAGGCTCAAGAAACAGAAATG 2340
2281 TTAAGAGACAATCCCATGATGCTGGCCCAATGCTAACACAGGCTCAAGAAACAGAAATG 2340
2341 TTAAGAGAGAGCATCCCTCGACCATCTGAATGAGAGATATGCTGTACCCCTTCCACCACA 2400
2341 TTAAGAGAGAGCATCCCTCGACCATCTGAATGAGAGATATGCTGTACCCCTTCCACCACA 2400
2401 AGTGGGACACCTCTGCATATCTGCTCCCTCTCTCTCTGTTAAGCCCGAGGAGGCCAT 2460
2401 AGTGGGACACCTCTGCATATCTGCTCCCTCTCTCTCTGTTAAGCCCGAGGAGGCCAT 2460
2461 CCACCCAGTGGTCTACAGACAGGCAATACACACACCAAGATAGCCTTCAGATCAAC 2520
2461 CCACCCAGTGGTCTACAGACAGGCAATACACACACCAAGATAGCCTTCAGATCAAC 2520
2521 ATGCATCACACTCAAGTGTAACTCTCAAGTGTCTTTTCTTTTCTCTGTTTCTTAT 2580
2521 ATGCATCACACTCAAGTGTAACTCTTTCAAGTGTCTTTTCTTTTCTCTGTTTCTTAT 2580
2581 TGTCTTCT 2640
2581 TGTCTTCT 2640
2641 GCCTAGAGCTAAATAATATAGAAATGATGTTATCTTGTGTGTGAGGAAAGGCCAGCT 2700
2641 GCCTAGAGCTAAATAATATAGAAATGATGTTATCTTGTGTGTGAGGAAAGGCCAGCT 2700
2701 GGCTTAAGTTCACACTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
2701 GGCTTAAGTTCACACTTTGTCTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
2761 GAAAGACACCTCTGANGCAGAGTCAGAGTCTGATCTACCCATCACTATTTTCTT 2820
2761 GAAAGACACCTCTGANGCAGAGTCAGAGTCTGATCTACCCATCACTATTTTCTT 2820
2821 CCATCATTTGCTTGGCTTGCCTCTCCACACCCGCTGACGTAATCGCATGGGAAG 2880
2821 CCATCATTTGCTTGGCTTGCCTCTCCACACCCGCTGACGTAATCGCATGGGAAG 2880
2881 CCAGGACAATGTTTGTGTTCTCTCTTGGGTAAAGGACTCCCTGAGCTCTGTGCTCT 2940
2881 CCAGGACAATGTTTGTGTTCTCTCTTGGGTAAAGGACTCCCTGAGCTCTGTGCTCT 2940
2941 CCAGTATGGTCCCTTTTCTCTTACAGATGCATATGTTTCTTCTTCTTCTTCTTCTTCT 3000
2941 CCAGTATGGTCCCTTTTCTCTTACAGATGCATATGTTTCTTCTTCTTCTTCTTCTTCT 3000
3001 ATTCTTAAATAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCAGACCTGC 3060
3001 ATTCTTAAATAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCAGACCTGC 3060
3061 ATTGTGTAGTGTGAATAGTGGATAAAGTGTGATGTCAAGAGAGTGAATCAACCTC 3120
3061 ATTGTGTAGTGTGAATAGTGGATAAAGTGTGATGTCAAGAGAGTGAATCAACCTC 3120
3121 TGCAAGACCAATCTTCTCTTCTTCTGAGTGTATTAAAGAAATACCTGAGTCTGTGTGTG 3180
3121 TGCAAGACCAATCTTCTCTTCTTCTGAGTGTATTAAAGAAATACCTGAGTCTGTGTGTG 3180
3181 TGTGTGTACCCAGTGTCAATCAATAAAGACCCAGTGTCAATGAGGAGGAGGAGGAGGAG 3240
3181 TGTGTGTACCCAGTGTCAATCAATAAAGACCCAGTGTCAATGAGGAGGAGGAGGAGGAGG 3240

QY 3241 AAAAAAAAAAAAAA 3255
DB 3241 AAAAAAAAAAAAAA 3255
RESULT 9
AX08977 3254 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 15 from Patent WO0116176.
DEFINITION AX08977
ACCESSION AX08977
VERSION AX08977.1 GI:13443971
KEYWORDS Mus musculus (house mouse).
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Liebert, R.
TITLE Transgenic animal model for neurodegenerative diseases
JOURNAL Patent: WO 0116176-A 15 28-MAR-2001;
Biofrontiers Pharmaceuticals GmbH (DE);
FEATURES
Location/Qualifiers
3241..3254
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 851 a 819 c 809 g 776 t
ORIGIN
Query Match 99.4%; Score 3243; DB 6; Length 3254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 0; Indels 1; Gaps 2;
QY 1 CTGAGCAGGCGAAGGGAGGAGGAGGCTGGATCACTAAACCTCAGACAGAAACGCTGGTGG 60
DB 1 CTGAGCAGGCGAAGGGAGGAGGAGGAGGCTGGATCACTAAACCTCAGACAGAAACGCTGGTGG 60
QY 61 AGGTCTGGGCGGCGCGAGTSCCGGCTAGGTCCTTCTGACCCGAGCCAGCCACCGCGCC 120
DB 61 AGGTCTGGGCGGCGCGAGTSCCGGCTAGGTCCTTCTGACCCGAGCCAGCCACCGCGCC 120
QY 121 CGTGACCATGATAGTGTGTTCTCAGATTCACCTCCAGCTATGCTTCCAGAGTGGAGTGG 180
DB 121 CGTGACCATGATAGTGTGTTCTCAGATTCACCTCCAGCTATGCTTCCAGAGTGGAGTGG 180
QY 181 ATTCTGACACACACATCTTCCAGCTCAAGGAAGTGGTGTCTAAGCAGACAGGGGGTCCAG 240
DB 181 ATTCTGACACACACATCTTCCAGCTCAAGGAAGTGGTGTCTAAGCAGACAGGGGGTCCAG 240
QY 241 CTGACCAAGCTGGTGTGATTTTTCGGGGAAGGAGCTTCGAAATCAGCTGACGTTCAAA 300
DB 241 CTGACCAAGCTGGTGTGATTTTTCGGGGAAGGAGCTTCGAAATCAGCTGACGTTCAAA 299
QY 301 ACTGTGACCTGGAAACACAGATATTGTACACATAGTACAGACACACGAGGAGAGTC 360
DB 301 ACTGTGACCTGGAAACACAGATATTGTACACATAGTACAGACACACGAGGAGAGTC 359
QY 361 ATGAACCAATATGATCTTGGAGGAGACGAACCCGAGACCTCAGAGGGCTCCATATGG 420
DB 361 ATGAACCAATATGATCTTGGAGGAGACGAACCCGAGACCTCAGAGGGCTCCATATGG 419
QY 421 AGTCCAGGAGCTTGACACAGAGTGGAGCTGAGCGGATACCTCTGGCGGTGGACTCTGG 480
DB 421 AGTCCAGGAGCTTGACACAGAGTGGAGCTGAGCGGATACCTCTGGCGGTGGACTCTGG 479
QY 481 GGTCTGGCGGTTCATCTGGACACACAGATTAAGAGGATTCAGAAAGCAGCCAGAGTCCAG 540
DB 481 GGTCTGGCGGTTCATCTGGACACACAGATTAAGAGGATTCAGAAAGCAGCCAGAGTCCAG 539
QY 541 TTAAGCCCACTCAACACAGCTTTTCTATCTAGTGTGAAGGCCCTTCCACAAAGGTCCAGC 600
DB 541 TTAAGCCCACTCAACACAGCTTTTCTATCTAGTGTGAAGGCCCTTCCACAAAGGTCCAGC 599

Db 1200 GCTCGGSGTTGTTTTCTSCGGGACTGTAAAGAGCATACCATGAAGGGGATTGGACT 1259
 Qy 1261 CACTGTCGAACCTCAGGAGCCACTTCTCAGGCCCTACAGGCTGGACAAAGAGCGCTG 1320
 Db 1260 CACTGTCGAACCTCAGGAGCCACTTCTCAGGCCCTACAGGCTGGACAAAGAGCGCTG 1319
 Qy 1321 AGCAAGCTCCTGGAGGAGGCTCCAAAGGAACCATCAGAGAGCCACCAAGCCTTGTC 1380
 Db 1320 AGCAAGCTCCTGGAGGAGGCTCCAAAGGAACCATCAGAGAGCCACCAAGCCTTGTC 1379
 Qy 1381 CTCGCTCAACGTGCCAATTGAAAAAAGAGGATGTATGCACATGAAGTGTCTCAAC 1440
 Db 1380 CTCGCTCAACGTGCCAATTGAAAAAAGAGGATGTATGCACATGAAGTGTCTCAAC 1439
 Qy 1441 CCCAGTGCAGCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1500
 Db 1440 CCCAGTGCAGCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1499
 Qy 1501 GAGATCACTGGTTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGACGCAAACTCAA 1560
 Db 1500 GAGATCACTGGTTGACGTGTAGAGAGAGATGTCACTTGGCCCTGGACGCAAACTCAA 1559
 Qy 1561 GGGAAACTCGAAGATTCTACCTTCTAGCCATTCTTCTCTCGATGCATATAGCA 1620
 Db 1560 GGGAAACTCGAAGATTCTACCTTCTAGCCATTCTTCTCTCGATGCATATAGCA 1619
 Qy 1621 CATAAATGCGCACACACAAACACAGGCTGCAGATTACAGAAAGAGCCCTAGATCTTTC 1680
 Db 1620 CATAAATGCGCACACACAAACACAGGCTGCAGATTACAGAAAGAGCCCTAGATCTTTC 1679
 Qy 1681 TAGGCAACCCACAGAAACACAGGCTGCAGATTACAGAAAGAGCCCTAGATCTTTC 1740
 Db 1680 TAGGCAACCCACAGAAACACAGGCTGCAGATTACAGAAAGAGCCCTAGATCTTTC 1739
 Qy 1741 CTCTGGCTCACTCAAAATGTAGAGCTTAGATGAGGTGCACTTGGTGGTGGTGGTGG 1800
 Db 1740 CTCTGGCTCACTCAAAATGTAGAGCTTAGATGAGGTGCACTTGGTGGTGGTGGTGG 1799
 Qy 1801 GAAGCCATGAGTGGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1860
 Db 1800 GAAGCCATGAGTGGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1859
 Qy 1861 CACTCTGAGGATCTCAGAAAGCACTTGAGATCTGAGAAAGCTGCTCATGTAGTA 1920
 Db 1860 CACTCTGAGGATCTCAGAAAGCACTTGAGATCTGAGAAAGCTGCTCATGTAGTA 1919
 Qy 1921 ATCATCTATTCCTCAAGGCGCCCTGAGTAGTCAAAATCTATTGTTTATCCCCCAAA 1980
 Db 1920 ATCATCTATTCCTCAAGGCGCCCTGAGTAGTCAAAATCTATTGTTTATCCCCCAAA 1979
 Qy 1981 CCTATCTTTCAAAATGGTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC 2040
 Db 1980 CCTATCTTTCAAAATGGTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC 2039
 Qy 2041 AACCAAGTGAACCTTAGAAAGCTTAATGGATGGCAGACTGCTTAATCGCAGGAGGA 2100
 Db 2040 AACCAAGTGAACCTTAGAAAGCTTAATGGATGGCAGACTGCTTAATCGCAGGAGGA 2099
 Qy 2101 CTCAGAGCCAAACCTTACCTCGTTCGTTTCTATCTGCAACTTTAGAAAGAAATGATC 2160
 Db 2100 CTCAGAGCCAAACCTTACCTCGTTCGTTTCTATCTGCAACTTTAGAAAGAAATGATC 2159
 Qy 2161 TTTTTCCTCCCTGAAAAGATAACAAAGTCTGCAATTTGGTTGGAGTATTCCTACTGCA 2220
 Db 2160 TTTTTCCTCCCTGAAAAGATAACAAAGTCTGCAATTTGGTTGGAGTATTCCTACTGCA 2219
 Qy 2221 GCCTGGAAGTTTAGCTTCACTGTGAATTTACAGAGAAAGTGCTATTAAGGGGGGGTTT 2280
 Db 2220 GCCTGGAAGTTTAGCTTCACTGTGAATTTACAGAGAAAGTGCTATTAAGGGGGGGTTT 2279
 Qy 2281 TTAAGAGCAATCCCATGATCTCGGCGAATGTAAACACAGGCTGCAAGAACCAATGT 2340
 Db 2280 TTAAGAGCAATCCCATGATCTCGGCGAATGTAAACACAGGCTGCAAGAACCAATGT 2339

Qy 2341 TTAAGAGCAATCCCATGATCTCGGCGAATGTAAACACAGGCTGCAAGAACCAATGT 2400
 Db 2340 TTAAGAGCAATCCCATGATCTCGGCGAATGTAAACACAGGCTGCAAGAACCAATGT 2399
 Qy 2401 AGTGGGAGCACTCTGSCATATCTGCTCCCTCTCTGCTGTTAAGCCCCAGGAGGCCCAT 2460
 Db 2400 AGTGGGAGCACTCTGSCATATCTGCTCCCTCTCTGCTGTTAAGCCCCAGGAGGCCCAT 2459
 Qy 2461 CCACCCAGTGGTCTACACAGAGGCAATACACACACAGAGATAGCTTCAGATCAAC 2520
 Db 2460 CCACCCAGTGGTCTACACAGAGGCAATACACACACAGAGATAGCTTCAGATCAAC 2519
 Qy 2521 ATGCATCACACCTCAAGTGTAACTTTCAAGGTTTCTTTCTTTTCTTTTCTTTTATT 2580
 Db 2520 ATGCATCACACCTCAAGTGTAACTTTCAAGGTTTCTTTCTTTTCTTTTCTTTTATT 2579
 Qy 2581 TGTTTTGGTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2640
 Db 2580 TGTTTTGGTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2639
 Qy 2641 GCTAGAGCTCAAAATCATACAGAAAGATGATGTTCTTGTGTGTGAGAAAGGTCAGCT 2700
 Db 2640 GCTAGAGCTCAAAATCATACAGAAAGATGATGTTCTTGTGTGTGAGAAAGGTCAGCT 2699
 Qy 2701 GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
 Db 2700 GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2759
 Qy 2761 GAAAGACACACTGTCAACAGAGCTCAGAGCTGTGATGTACCCATCACTATTTTTTT 2820
 Db 2760 GAAAGACACACTGTCAACAGAGCTCAGAGCTGTGATGTACCCATCACTATTTTTTT 2819
 Qy 2821 CCAATCTGTGTGCTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 2880
 Db 2820 CCAATCTGTGTGCTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 2879
 Qy 2881 CAGGCAATGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 2940
 Db 2880 CAGGCAATGTTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 2939
 Qy 2941 CCAAGTATGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3000
 Db 2940 CCAAGTATGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2999
 Qy 3001 ATTCTTAAATTAACCCAAAGACAGGCTTCCACAGCTGTGAGCATGATATCAGCTTCT 3060
 Db 3000 ATTCTTAAATTAACCCAAAGACAGGCTTCCACAGCTGTGAGCATGATATCAGCTTCT 3059
 Qy 3061 ATTCTTAAATTAACCCAAAGACAGGCTTCCACAGCTGTGAGCATGATATCAGCTTCT 3120
 Db 3060 ATTCTTAAATTAACCCAAAGACAGGCTTCCACAGCTGTGAGCATGATATCAGCTTCT 3119
 Qy 3121 TGCAGAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
 Db 3120 TGCAGAGCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3179
 Qy 3181 TGTGTGTACCACTGTGATCAATTAAGAGCCAGGCTGTCAATGAAGAAAGAAAGAA 3240
 Db 3180 TGTGTGTACCACTGTGATCAATTAAGAGCCAGGCTGTCAATGAAGAAAGAAAGAA 3239
 Qy 3241 AAAAAAAAAAAAAA 3255
 Db 3240 AAAAAAAAAAAAAA 3254

RESULT 1:
 LOCUS AX089876 3253 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 14 from Patent WO01/6176
 ACCESSION AX089876
 VERSION AX089876.1 GI:3443970
 KEYWORDS

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.	
REFERENCE	Juebert,H.
AUTHORS	Transgenic animal model for neurodegenerative diseases
TITLE	Patent: WO 016176-A 14 09-MAR-2001;
JOURNAL	Biofrontiera Pharmaceuticals GmbH (DE)
FEATURES	Location/Qualifiers
source	i..3253
/organism="Mus musculus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10090"	
BASE COUNT	851 a _818 c 808 g 776 t
ORIGIN	
Query March	99.8%; Score 3241; DB 6; Length 3253;
Best local Similarity	99.9%; Pred. No. 3;
Matches 3253; Conservative 0; Mismatches 0; Indels 2; Gaps 1;	
QY	1 CTGAGCGGGGAAGGGGAGGAGGCCTGGATCACTAAACCTGACAGAAACGCTGTGGG 60
DB	1 CTGAGCGGGGAAGGGGAGGAGGCCTGGATCACTAAACCTGACAGAAACGCTGTGGG 60
QY	61 AGCCTGGGCGGGCGCCAGTGCCTGGCTAGTTCCTTGACGCCGACGCCACCACCCC 120
DB	61 AGCCTGGGCGGGCGCCAGTGCCTGGCTAGTTCCTTGACGCCGACGCCACCACCCC 120
QY	121 CGGTGACCATGATGTTGTTGTGAGGTTCAACTCGAGCTATGCTTCCAGTGGAGTGC 180
DB	121 CGGTGACCATGATGTTGTTGTGAGGTTCAACTCGAGCTATGCTTCCAGTGGAGTGC 180
QY	181 ACTGTGACACCATGATCTTSCAGCTCAAGGAGTGGTCTTAAGCGAGCGGGGTTCCAG 240
DB	181 ATTCTGACACCATGATCTTSCAGCTCAAGGAGTGGTCTTAAGCGAC--GGGTTCCAG 238
QY	241 CTGACCATGCTGGTGTGATTTTTCCGGGAAGAGGCTTCCGAATCACCTGACGTTCAA 300
DB	239 CTGACCATGCTGGTGTGATTTTTCCGGGAAGAGGCTTCCGAATCACCTGACGTTCAA 298
QY	301 ACTGTGACCTGGAAACAACAGAGATTGTTACACATAGTACAGAGACACGGAGGAAGTC 360
DB	299 ACTGTGACCTGGAAACAACAGAGATTGTTACACATAGTACAGAGACACGGAGGAAGTC 358
QY	361 ATGAACAATAATGATCTGGAGGGGAGCAACCCAGAGACCTCAGAGGCTCCAATGCG 420
DB	359 ATGAACAATAATGATCTGGAGGGGAGCAACCCAGAGACCTCAGAGGCTCCAATGCG 418
QY	421 AGTCCAGGAGCTTGACACGAGTGGACCTGAGCAGCCATACCTCGGGTGGACTCTGG 480
DB	419 AGTCCAGGAGCTTGACACGAGTGGACCTGAGCAGCCATACCTCGGGTGGACTCTGG 479
QY	481 GGCTGGCGGTGATCTGGACACAGAGATTAGAGGATTCAGAGACGACGAGGATCCAG 540
DB	479 GGCTGGCGGTGATCTGGACACAGAGATTAGAGGATTCAGAGACGACGAGGATCCAG 538
QY	541 TTAACCCACCTACACAGCTTTTATCTACTGCAAGGCCCTGCCACAAGGTTCCAGC 600
DB	539 TTAACCCACCTACACAGCTTTTATCTACTGCAAGGCCCTGCCACAAGGTTCCAGC 598
QY	601 CTGAAAAGCTCCGAGTTTCACTGTGGACCTGCAAAACAAGCAACCTTCACTTGGCCCC 660
DB	599 CTGAAAAGCTCCGAGTTTCACTGTGGACCTGCAAAACAAGCAACCTTCACTTGGCCCC 658
QY	661 GCCCATCTTGTGGGACGATGCTTAAATTCGAACCCGATGAGTGGTGGAGTCCAGTCTC 720
DB	659 GCCCATCTTGTGGGACGATGCTTAAATTCGAACCCGATGAGTGGTGGAGTCCAGTCTC 718
QY	721 CAGACTGCCCTGGAAACAGAGCTGAATTTTCTTTAATGTGGACACACACCAACCTCAG 780
DB	719 CAGACTGCCCTGGAAACAGAGCTGAATTTTCTTTAATGTGGACACACACCAACCTCAG 778

[illegible]

Db 301 ACTGTGACCTGGAAACAACAGAGTATTGTACACATAGTACAGAGACCAG--GGAGAGTC 358
QY 361 ATGAACAATAATGCACTCTGGAGGGGACGAACCCAGAGACACTCAGAGGGGTCCATATGGG 423
Db 359 ATGAACAATAATGCACTCTGGAGGGGACGAACCCAGAGACACTCAGAGGGGTCCATATGGG 418
QY 421 AGTCAGAGAGCTTACACAGAGTGAACCTGAGACGCCATACCTGCGCGTGGATCTCTGTG 490
Db 419 AGTCAGAGAGCTTACACAGAGTGAACCTGAGACGCCATACCTGCGCGTGGACTCTGTG 478
QY 481 GGCTGGGGCTCATCTGGACACACACAGTAAAGGGGATTGAGAGGACGACAGAGGTCCAG 540
Db 479 GGCCTGGCGGTCACTCTGGACACACACAGTAAAGGGGATTGAGAGGACGACAGAGGTCCAG 538
QY 541 TTAACCCACCTACAACAGCTTTTTCATCTACTCCAAAGGCCCTTGCCACAAAGTCCAGC 600
Db 539 TTAACCCACCTACAACAGCTTTTTCATCTACTCCAAAGGCCCTTGCCACAAAGTCCAGC 598
QY 601 CTGGAAGAGCTCCGAGTTCAAGTGTGGACCTGCGCAACAAAGCAACCTCACCTTGGCCAGG 660
Db 599 CTGGAAGAGCTCCGAGTTCAAGTGTGGACCTGCGCAACAAAGCAACCTCACCTTGGCCAGG 658
QY 661 GCCCATCTGCTGGGACGATGCTTAATTCCAAACCGGATGAGTGGTGAAGTGCAGTCTC 720
Db 659 GCCCATCTGCTGGGACGATGCTTAATTCCAAACCGGATGAGTGGTGAAGTGCAGTCTC 718
QY 721 CAGACTGCCCTGGAACAGAGCTGAATTTTCTTTAAATGTGAGACACACCCAACTCAG 780
Db 719 CAGACTGCCCTGGAACAGAGCTGAATTTTCTTTAAATGTGAGACACACCCAACTCAG 778
QY 781 ACAAGGACAGTCCGTAGCTTTGAACCTGATCACACAGCAGCGGAGGATCCCTTSCA 840
Db 779 ACAAGGACAGTCCGTAGCTTTGAACCTGATCACACAGCAGCGGAGGATCCCTTSCA 838
QY 841 TAGCGTGACAGATGTCCAGAGCCCTGTCCGTGCTTCCAGTGTAAACACCGCTCAAGTGA 900
Db 839 TAGCGTGACAGATGTCCAGAGCCCTGTCCGTGCTTCCAGTGTAAACACCGCTCAAGTGA 898
QY 901 TCTGTTTGGACTGTTTCCACTTGATTTGTGTCAACAGACTCAAGATCGGAGTTGTTC 960
Db 899 TCTGTTTGGACTGTTTCCACTTGATTTGTGTCAACAGACTCAAGATCGGAGTTGTTC 958
QY 961 ACAGTGTCCACTTGGCTACTCCCTGGCGTGTAGTGGCTGCCCACTCCCTGATTA 1020
Db 959 ACAGTGTCCACTTGGCTACTCCCTGGCGTGTAGTGGCTGCCCACTCCCTGATTA 1018
QY 1021 AAGAGCTCCATCACCTTCAAGATCTCTGGAGAGAGCAGTACACTAGGTACAGCAGTATG 1080
Db 1019 AAGAGCTCCATCACCTTCAAGATCTCTGGAGAGAGCAGTACACTAGGTACAGCAGTATG 1078
QY 1081 GGGCGAGGAATGGTCTGCAAAATGGAGAGTGTCTGTGCCCGCTGCTGGTGTGGAG 1140
Db 1079 GGGCGAGGAATGGTCTGCAAAATGGAGAGTGTCTGTGCCCGCTGCTGGTGTGGAG 1138
QY 1141 CTGGAAGTCTACCTGAAACAGGCGCCAGAGAAAGTCACTCTGGAAGGGGCAACGGCCCTGG 1200
Db 1139 CTGGAAGTCTACCTGAAACAGGCGCCAGAGAAAGTCACTCTGGAAGGGGCAACGGCCCTGG 1198
QY 1201 GCTGGCGGTGTGTTTCTGCGGAGCTCTAAGAGAGCATCATGAAGGGGATTTGCAACT 1260
Db 1199 GCTGGCGGTGTGTTTCTGCGGAGCTCTAAGAGAGCATCATGAAGGGGATTTGCAACT 1258
QY 1261 CACTGCTGAAACCTCAGAGCCACTTCTCAGGCTACAGAGGTGGAACAAAGAGCCGCTG 1320
Db 1259 CACTGCTGAAACCTCAGAGCCACTTCTCAGGCTACAGAGGTGGAACAAAGAGCCGCTG 1318
QY 1321 AGCAAGCTGCTGGAGAGGAGCTTCAAAGGAAACCATCAAGAAACACCAACAGCCCTTGT 1380
Db 1319 AGCAAGCTGCTGGAGAGGAGCTTCAAAGGAAACCATCAAGAAACACCAACAGCCCTTGT 1378
QY 1381 CTGCGTCAACCTGCCAATTTGAAAAAAGAGGAGGATGTATCCATGAAGTGTCTCAGC 1440
Db 1379 CTGCGTCAACCTGCCAATTTGAAAAAAGAGGAGGATGTATCCATGAAGTGTCTCAGC 1438

QY 1441 CCCAGTGCACAGCTGGAGTGTGCTGGAACCTGTGGCTGTGAGTGSAAACCGAGCCTGCATGG 1500
Db 1439 CCCAGTGCACAGCTGGAGTGTGCTGGAACCTGTGGCTGTGAGTGSAAACCGAGCCTGCATGG 1498
QY 1501 GAGATCACTGTTTGAAGTGTAGAGAGAGATGTCACTTGGCCCTGGAGCAGACAACTCAA 1560
Db 1499 GAGATCACTGTTTGAAGTGTAGAGAGAGATGTCACTTGGCCCTGGAGCAGACAACTCAA 1558
QY 1561 GGGAAATCCCAAGATTTCTACTTCCCTTAGCCATTTCTTCTCTCCATGCATATAGA 1620
Db 1559 GGGAAATCCCAAGATTTCTACTTCCCTTAGCCATTTCTTCTCTCCATGCATATAGA 1618
QY 1621 CATAAATGGCAGACACACAAACACAGAGTGCAGATTACAGAGAGAGCCCTTAGATCTTTC 1680
Db 1619 CATAAATGGCAGACACACAAACACAGAGTGCAGATTACAGAGAGAGCCCTTAGATCTTTC 1678
QY 1681 TAGGGCACCCACAGANNAACACACAGACCCGCTGGCCAGGGGGAGGAGGACACTTCAGC 1740
Db 1679 TAGGGCACCCACAGANNAACACACAGACCCGCTGGCCAGGGGGAGGAGGACACTTCAGC 1738
QY 1741 CTCTGGCTCACTCGAATGTCCAGAGCTTAGATGAGGTGCACCTTTGGTTTGGATTCGTA 1800
Db 1739 CTCTGGCTCACTCGAATGTCCAGAGCTTAGATGAGGTGCACCTTTGGTTTGGATTCGTA 1798
QY 1801 GAAGCCATGAGTGGGTGGGAAAGTGTTCAGAGCTTTTCCAGCGTTGTTCACACGCGCTGGTAACTAA 1860
Db 1799 GAAGCCATGAGTGGGTGGGAAAGTGTTCAGAGCTTTTCCAGCGTTGTTCACACGCGCTGGTAACTAA 1858
QY 1861 CACCTCTCAGGATTTCTCAGAGGACACTTTGAGATCTCAGAGAGCCCTGCTCATCTAGTA 1920
Db 1859 CACCTCTCAGGATTTCTCAGAGGACACTTTGAGATCTCAGAGAGCCCTGCTCATCTAGTA 1918
QY 1921 ATCATCTATTCCTCAAGAGGCGCTTGCAGTAGTCAAAACTATTTTATCCCGCCCAAT 1980
Db 1919 ATCATCTATTCCTCAAGAGGCGCTTGCAGTAGTCAAAACTATTTTATCCCGCCCAAT 1978
QY 1981 CCTATCTTCAAAATGGTGTGATGAGATTACACCCCTCTGTGTACTTAATTCAGCTTATC 2040
Db 1979 CCTATCTTCAAAATGGTGTGATGAGATTACACCCCTCTGTGTACTTAATTCAGCTTATC 2038
QY 2041 AACCAAGTGAGAACTCAGGAAAGCTAATTCGATGGCAGACTGCTTAAATGCGAGGAGGA 2100
Db 2039 AACCAAGTGAGAACTCAGGAAAGCTAATTCGATGGCAGACTGCTTAAATGCGAGGAGGA 2098
QY 2101 CTGAAGAGCCAAAGCTACTTCCGTTTCTATCTGCAACTTTAGAGAAATGATC 2160
Db 2099 CTGAAGAGCCAAAGCTACTTCCGTTTCTATCTGCAACTTTAGAGAAATGATC 2158
QY 2161 TTTTTCCTCCCTGAAAGATACAAAGTGTGCAATTTGGTTTGGAGTATTCCTACTGCA 2220
Db 2159 TTTTTCCTCCCTGAAAGATACAAAGTGTGCAATTTGGTTTGGAGTATTCCTACTGCA 2218
QY 2221 CCTGGAAGTTTACCTTCAGTGTAAATTTAAACAGAGAAAGTGCCTATAAAGGGGCGTTT 2280
Db 2219 CCTGGAAGTTTACCTTCAGTGTAAATTTAAACAGAGAAAGTGCCTATAAAGGGGCGTTT 2278
QY 2281 TTAAGAGACATCCGATCATCTGCGGCATCTGCAACAGAGGCTCAAGAAACCAATGT 2340
Db 2279 TTAAGAGACATCCGATCATCTGCGGCATCTGCAACAGAGGCTCAAGAAACCAATGT 2338
QY 2341 TTAAGAGAGGAGCTCCCTGACCATCTGAATGAGATGCTGTGACCCCTTCCACACA 2400
Db 2339 TTAAGAGAGGAGCTCCCTGACCATCTGATGATGAGATGCTGTGACCCCTTCCACACA 2398
QY 2401 AGTGGGAGACCTTGCATATCTGCTCCCTCTCTGTGTGTAAAGCCGAGGAGCCGCAAT 2460
Db 2399 AGTGGGAGACCTTGCATATCTGCTCCCTCTCTGTGTGTAAAGCCGAGGAGCCGCAAT 2458
QY 2461 CCACCCAGTGTCTTACAGAGGCGCATACACACACACCAAGATAGCTTTCAGATCAAC 2520
Db 2459 CCACCCAGTGTCTTACAGAGGCGCATACACACACCAAGATAGCTTTCAGATCAAC 2518

QY	3181	TGTTGGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATGTAAGAAAAA	3240
QY <td>3179</td> <td>TGTTGGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATGTAAGAAAAA</td> <td>3238</td>	3179	TGTTGGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATGTAAGAAAAA	3238
QY	3241	AAAAA	3255
Db	3239	AAAAA	3253
RESULT 14	AX136843	Sequence 16 from Patent EP1081255.	3253 bp DNA linear PAT 30-MAY-2001
LOCUS	DEFINITION	AX136843	
ACCESSION	VERSION	AX136843.1	GI:14273208
KEYWORDS	SOURCE	Mus musculus (house mouse)	
ORGANISM	REFERENCE	1	
AUTHORS	TITLE	Transgenic animal model for neurodegenerative diseases	
JOURNAL	Patent:	EP 1081255-A 16 07-MAR-2001;	
FEATURES	Location/Qualifiers	1..3253	
source	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
BASE COUNT	851 a	818 c	808 g 776 t
ORIGIN	Query March	99.6%	Score 3241; DB 6; Length 3253;
	Best Local Similarity	99.9%	Pred. No. C;
	Matches 3253; Conservative	0;	Mismatches 0; Indels 2; Gaps 1;
QY	1	CTCAGGAGGGGAAGGGGAGGAGGAGGCGCTGATGACTAAACCTGACAGAAACGCTGTGGG	60
Db	1	CTCAGGAGGGGAAGGGGAGGAGGAGGCGCTGATGACTAAACCTGACAGAAACGCTGTGGG	60
QY	61	AGGCTCGGGCGGGCGCCAGTGCCTGAGTGCCTCTCGACCGGAGGACACACCGCC	120
Db	61	AGGCTCGGGCGGGCGCCAGTGCCTGAGTGCCTCTCGACCGGAGGACACACCGCC	120
QY	121	CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGC	180
Db	121	CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGC	180
QY	181	ATTCTGACACCAAGCATTTTCAGCTCAAGCAAGTGGTTGCTTAAGCGACAGGGGGTCCAG	240
Db	181	ATTCTGACACCAAGCATTTTCAGCTCAAGCAAGTGGTTGCTTAAGCGACAGGGGGTCCAG	240
QY	241	CTGACCAAGCTGGTGTGATTTTCGGGGAGAGAGTTCCTCGATCACTGAGGTTCAA	300
Db	241	CTGACCAAGCTGGTGTGATTTTCGGGGAGAGAGTTCCTCGATCACTGAGGTTCAA	300
QY	301	ACTGTGACCTGGAAACACAGAGATTTGTACACATAGTACAGAGACACCGAGAGAGTCC	360
Db	301	ACTGTGACCTGGAAACACAGAGATTTGTACACATAGTACAGAGACACCG--GGAGAGTCC	358
QY	361	ATGAACCAATGCTATCTGGAGGGAGCAACCCAGAGACCTCAGAGGGCTCCATATGGG	420
Db	359	ATGAACCAATGCTATCTGGAGGGAGCAACCCAGAGACCTCAGAGGGCTCCATATGGG	418
QY	421	AGTCCAGGAGCTTGACACAGTGGACCTTGAGCAGCCATACCTGCCGGTGGACTCTGTGG	480
Db	419	AGTCCAGGAGCTTGACACAGTGGACCTTGAGCAGCCATACCTGCCGGTGGACTCTGTGG	478
QY	481	GGCTGGCGGTCTATCTTGGGACACAGACAGTAAAGAGGATTCAGAGCAGCCAGAGTCCAG	540
Db	479	GGCTGGCGGTCTATCTTGGGACACAGACAGTAAAGAGGATTCAGAGCAGCCAGAGTCCAG	538

```
QY 1621 CATAAATGGCACACACAAACACAGGCTGCAGATTACAGAAAGAGCCCTTAGATCTCTTC 1680
Db 1619 CATAAATGGCACACACAAACACAGGCTGCAGATTACAGAAAGAGCCCTTAGATCTCTTC 1678
QY 1681 TAGGACCCACAGAAACACACACACCCGCTGCCCGAGGGGAGGAGGACCTTCAGC 1740
Db 1679 TAGGACCCACAGAAACACACACACCCGCTGCCCGAGGGGAGGAGGACCTTCAGC 1738
QY 1741 CTCTGGCTCACTCGAATGTCAAGAGCTTAGAAGAGGGTGCACTTTGGTTGGATTCTGTA 1800
Db 1739 CTCTGGCTCACTCGAATGTCAAGAGCTTAGAAGAGGGTGCACTTTGGTTGGATTCTGTA 1798
QY 1801 GAAGCCATGAGTGAGGTGGGAAGTGTTCAGAGGTGTTCAGAGGCTTCAGAGGCTGGGTAAGTAA 1860
Db 1799 GAAGCCATGAGTGAGGTGGGAAGTGTTCAGAGGTGTTCAGAGGCTTCAGAGGCTGGGTAAGTAA 1858
QY 1861 CACCTCTGAGGATCTCAGAGGACACTTGAGATCTGAGGAACCGTCTCTCATGTAGTA 1920
Db 1859 CACCTCTGAGGATCTCAGAGGACACTTGAGATCTGAGGAACCGTCTCTCATGTAGTA 1918
QY 1921 ATCATCTATTCCCAAGGCGCCCTGCAGTAGTCAAAACTATTGTCTATCCGCCCAAT 1980
Db 1919 ATCATCTATTCCCAAGGCGCCCTGCAGTAGTCAAAACTATTGTCTATCCGCCCAAT 1978
QY 1981 CCTATCTTTACAAATGCTGTATGAGATTACACCCCTCTGTGTACTTAATCAGCTTATC 2040
Db 1979 CCTATCTTTACAAATGCTGTATGAGATTACACCCCTCTGTGTACTTAATCAGCTTATC 2038
QY 2041 AACCAAGTGAGAACTAGGAAGCTAATTTGATGCGAGACTGCTTAATTCGACGGAGGA 2100
Db 2039 AACCAAGTGAGAACTAGGAAGCTAATTTGATGCGAGACTGCTTAATTCGACGGAGGA 2098
QY 2101 CTCAGAACCAAACTACTTCGGTTCGTTTCATTATCTCAACTTTAGAAAGAAATGATC 2160
Db 2099 CTCAGAACCAAACTACTTCGGTTCGTTTCATTATCTCAACTTTAGAAAGAAATGATC 2158
QY 2161 TTTTTCCTGCTGAAAGATAACAAAGCTGCAATTTGGTTGGAGTATTCCTACTGCA 2220
Db 2159 TTTTTCCTGCTGAAAGATAACAAAGCTGCAATTTGGTTGGAGTATTCCTACTGCA 2218
QY 2221 GCTCGAAGTTTAGCTTCACCTGGAATTTACAGAGAAGTGCCCTATAAAGGGGGCTTT 2280
Db 2219 GCTCGAAGTTTAGCTTCACCTGGAATTTACAGAGAAGTGCCCTATAAAGGGGGCTTT 2278
QY 2281 TTAAGAGACAATCCCATGATGCTGGCCCAATGCTTAACAACAGGGTCAAGAAACAATGT 2340
Db 2279 TTAAGAGACAATCCCATGATGCTGGCCCAATGCTTAACAACAGGGTCAAGAAACAATGT 2338
QY 2341 TTATAGAGGAGCATCCCTCGACCATCTGAATGAGATGATGCTGACCCCTTCACCACA 2400
Db 2339 TTATAGAGGAGCATCCCTCGACCATCTGAATGAGATGATGCTGACCCCTTCACCACA 2398
QY 2401 AGTGGGGACACCTCTGCATATCTGCTCCTCTCTGCTGTTAAGCCCGGAGGAGCCCAT 2460
Db 2399 AGTGGGGACACCTCTGCATATCTGCTCCTCTCTGCTGTTAAGCCCGGAGGAGCCCAT 2458
QY 2461 CCACCCAGTGTCTACAGACAGGCAATACACACACCAAGATAGACCTTCAGATCAAC 2520
Db 2459 CCACCCAGTGTCTACAGACAGGCAATACACACACCAAGATAGACCTTCAGATCAAC 2518
QY 2521 ATGCATCAGCTCAAGTGTTAATCTTTCAGGTTTCTCTTTCTCTCTCTCTCTCTCTCT 2580
Db 2519 ATGCATCAGCTCAAGTGTTAATCTTTCAGGTTTCTCTTTCTCTCTCTCTCTCTCTCT 2578
QY 2581 TGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
Db 2579 TGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2638
QY 2641 GCTAGAGCTAAAAATCATATAGAAATGATGTATCTTTGTGTGTGTGTGTGTGTGTGTGT 2700
Db 2639 GCTAGAGCTAAAAATCATATAGAAATGATGTATCTTTGTGTGTGTGTGTGTGTGTGTGT 2698
QY 2701 GGCCTAAGTTACACTTTTGTGCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
```

```
Db 2699 GGCCTAAGTTACACTTTTGTGCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2758
QY 2761 GAAGAGACACCTGTCAAGACAGCAGTCAAGAGTCTGAGTGCACCCATCACTATTTT 2820
Db 2759 GAAGAGACACCTGTCAAGACAGCAGTCAAGAGTCTGAGTGCACCCATCACTATTTT 2818
QY 2821 CCATCATTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2880
Db 2819 CCATCATTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2878
QY 2881 CAGAGCAATTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2940
Db 2879 CAGAGCAATTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2938
QY 2941 CAGATATGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 3000
Db 2939 CAGATATGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2998
QY 3001 ATTCTTAAATTAACCCAAAGACAGGCAATCCACATTTGTTGAGCATGATCAAGACCTTC 3060
Db 2999 ATTCTTAAATTAACCCAAAGACAGGCAATCCACATTTGTTGAGCATGATCAAGACCTTC 3058
QY 3061 ATTGTTGAGTGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3120
Db 3059 ATTGTTGAGTGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3118
QY 3121 TGCAAGCAATCTTTCTCTTTCTGTTGAGTGTATTAAAGAAATACCTGAGTGTGTGTG 3180
Db 3119 TGCAAGCAATCTTTCTCTTTCTGTTGAGTGTATTAAAGAAATACCTGAGTGTGTGTG 3178
QY 3181 TGTGTGATCCAGAGTGTCAATCAATTAAGACCCAGAGCTGTCAATGAAAAA 3240
Db 3179 TGTGTGATCCAGAGTGTCAATCAATTAAGACCCAGAGCTGTCAATGAAAAA 3238
QY 3241 AAAAAA 3255
Db 3239 AAAAAA 3253
```

RESULT 15
AF250293
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
RECEIVED
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AF250293
Mus musculus parkin mRNA, complete cds, alternatively spliced.
AF250293.1 GI:10179808
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathia; Muridae; Murinae; Mus;
Stichel, C.C., Augustin, M., Kuhn, K., Zhu, X.-R., Engels, P., Umer, C.
and Lubbert, H.
Parkin expression in the adult mouse brain
Unpublished
2 (bases 1 to 3226)
Engels, P., Augustin, M., Stichel, C.C. and Lubbert, H.
Direct Submission
Submitted (29-MAR-2003) Biofrontera Pharmaceuticals GmbH,
Hemelrathweg 201, Leverkusen 51377, Germany
Location/Qualifiers
1. 3226
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BA7.2/c"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_type="kidney"
/note="alternatively spliced"
/codon_start=1
/product="parkin"

/protein_id="AAG13890.1"									
/db_xref="GI:10179809"									
/translation="MIVFVRNFSYGFPEVDSSTLILQKXWAKROGVADQLRVI									
FAGKELPNHLITVQNCLEQSIIVH:VORPRRSHEINASGGDEPOSTSEGISWESRSL									
TRVDSLSTPLVDVSLVILDTDSKRDSEARAPVKPYNSFFIYICGPCHKVQPGK									
LRVQGTCKQATLTLAAGPSCDVLI PNRMSECCSPDCPGTRAEFFFKCGAPHTSD									
KDTSLVNLITNSRRSIPICIACTVRSPLVFCNHRHVICUDCFHLFYCVTRLNDRQF									
VHDAQGLYSLPCVAGPNSLILKELHFR:LGBEQVTRYQYGAEBECVLMGSLVCLPRP									
GCGAGLLPQGGKRVKCEGNGLIGCFVCRCKEAYHEGDCDSLLESGATSOAYRV									
DKXAEQARWEERASKETIKKTKPCPRKNVPIEKNGGVCHMKCPQPOCKLEKNKCGC									
EWNRAKCGDHFVDV"									
BASE COUNT	846 a	813 c	793 g	774 t					
ORIGIN									
Query Match	99.1%; Score 3226; DB 10; Length 3226;								
Best Local Similarity	100.0%; Pred. No. 0;								
Matches 3226;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	30	GATGACTAAACCTGACAGAAACCTGGTGGAGGCTCGGGCGGGGCCAGTGC	89						
DB	1	GATGACTAAACCTGACAGAAACCTGGTGGAGGCTCGGGCGGGGCCAGTGC	50						
QY	90	GGTCTCTTCGACCGCAGCCACACCCCGCTGACCATGATGCTGTCAGGTTG	149						
DB	61	GGTCTCTTCGACCGCAGCCACACCCCGCTGACCATGATGCTGTCAGGTTG	120						
QY	150	AAGTCCAGCTATGGCTTCCAGTGGAGTGGATTCTGACACAGCATCTTGCA	209						
DB	121	AAGTCCAGCTATGGCTTCCAGTGGAGTGGATTCTGACACAGCATCTTGCA	180						
QY	210	GAAGTGTTGCTTAAGCAGACAGGGGGTTCAGCTGACAGCTGGGTGATTTT	269						
DB	191	GAAGTGTTGCTTAAGCAGACAGGGGGTTCAGCTGACAGCTGGGTGATTTT	240						
QY	270	AAGGAGCTTCGGAATCACCTGACGGTTCGAACTGTGACCTGGAACAGAG	329						
DB	241	AAGGAGCTTCGGAATCACCTGACGGTTCGAACTGTGACCTGGAACAGAG	300						
QY	330	CACATAGTACAGAGACCCAGGAGAGAGTCTATGAACAAATGATCTGAG	389						
DB	301	CACATAGTACAGAGACCCAGGAGAGAGTCTATGAACAAATGATCTGAG	360						
QY	390	CCGAGAGACCTCAGAGGCTTCATATGGAGTGCAGAGCTTGACAGAGTGG	449						
DB	361	CCGAGAGACCTCAGAGGCTTCATATGGAGTGCAGAGCTTGACAGAGTGG	420						
QY	450	AGCAGCATACCTTCGCGTGGACTCTGTGGGCTGGCGTCACTTCTGACAC	509						
DB	421	AGCAGCATACCTTCGCGTGGACTCTGTGGGCTGGCGTCACTTCTGACAC	480						
QY	510	AAGAGGATTCAGAGCAGCAGAGTTCAGTTAAACCCACCTTACAGAGCTT	569						
DB	481	AAGAGGATTCAGAGCAGCAGAGTTCAGTTAAACCCACCTTACAGAGCTT	540						
QY	570	TACTGCAAGAGCCCTGCCACAGAGTCCAGCTCGAAAGCTCCGAGTTGAG	629						
DB	541	TACTGCAAGAGCCCTGCCACAGAGTCCAGCTCGAAAGCTCCGAGTTGAG	600						
QY	630	TGCAACCAAGCAACCTTACCTTGGCCAGGSCCATCTTGTGGGACGATGT	689						
DB	601	TGCAACCAAGCAACCTTACCTTGGCCAGGSCCATCTTGTGGGACGATGT	660						
QY	690	CCAAACCGGATGAGTGTGAGTGCAGTCCAGACTGCCCTGGAAACAGAG	749						
DB	661	CCAAACCGGATGAGTGTGAGTGCAGTCCAGACTGCCCTGGAAACAGAG	720						
QY	750	TTCTTTAAATGTGGAGCACACCCAACTCAGACAGGACAGCTGGTAGCTT	809						
DB	721	TTCTTTAAATGTGGAGCACACCCAACTCAGACAGGACAGCTGGTAGCTT	780						
QY	810	ATCACAGCAACAGGCGCAGCATCCCTTGCATAGCGTGACAGATGTGAGG	869						
DB	781	ATCACAGCAACAGGCGCAGCATCCCTTGCATAGCGTGACAGATGTGAGG	840						

Qy	970	CTGCTCTTCCAGTGTAAACCACCGTCACGTGATCTGTTGGACTGTTTCCACT	929						
Ds	841	CTGCTCTTCCAGTGTAAACCACCGTCACGTGATCTGTTGGACTGTTTCCACT	900						
Qy	930	GTCACAGACTCAACGATCGGAGTGTGTCCAGATGCTCAACTTGGCTACT	989						
Ds	901	GTCACAGACTCAACGATCGGAGTGTGTCCAGATGCTCAACTTGGCTACT	960						
Qy	990	TGTGTAGCTGGCTGCCAACTCCCTGATTAAAGAGCTCCACTCAGATGCT	1049						
Ds	961	TGTGTAGCTGGCTGCCAACTCCCTGATTAAAGAGCTCCACTCAGATGCT	1020						
Qy	1050	GAAGAGCAGTACACTAGTATACCAACAGATAGGGCCGAGGAATGCGTGT	1109						
Ds	1021	GAAGAGCAGTACACTAGTATACCAACAGATAGGGCCGAGGAATGCGTGT	1080						
Qy	1110	GGTGTGTGTGCCCGCTCTGTGTGTGGAGCTTGGACTGTCTGACAGAGG	1169						
Ds	1081	GGTGTGTGTGCCCGCTCTGTGTGTGGAGCTTGGACTGTCTGACAGAGG	1140						
Qy	1170	AAAGTCACTTGCAGAGGGGCAACGGGCTGGGCTGAGGGTGTCTCTCTCT	1229						
Ds	1141	AAAGTCACTTGCAGAGGGGCAACGGGCTGGGCTGAGGGTGTCTCTCTCT	1200						
Qy	1230	AAGGAACATACCATATGAAGGGATATGCAATGACTGTGCAACCTCTCAG	1289						
Ds	1201	AAGGAACATACCATATGAAGGGATATGCAATGACTGTGCAACCTCTCAG	1260						
Qy	1290	CAGGCTACAGGCTGAGCAAAAGAGCGCTGAGCAAGCTCGCTGGAGAGG	1349						
Ds	1261	CAGGCTACAGGCTGAGCAAAAGAGCGCTGAGCAAGCTCGCTGGAGAGG	1320						
Qy	1350	GAAGACCATCAAGAGAGCAACCAAGCTTGTCTGTGTGCAACCTGCAAA	1409						
Ds	1321	GAAGACCATCAAGAGAGCAACCAAGCTTGTCTGTGTGCAACCTGCAAA	1380						
Qy	1410	GGAGGATGTATGACATGAAAGTGTCTCAGCGCCAGTGCMAAGTGGAGT	1469						
Ds	1381	GGAGGATGTATGACATGAAAGTGTCTCAGCGCCAGTGCMAAGTGGAGT	1440						
Qy	1470	TGTGGCTGTGATGGAACCGAGCTGCATGGGAGTCACTGCTTGGAGTGT	1529						
Ds	1441	TGTGGCTGTGATGGAACCGAGCTGCATGGGAGTCACTGCTTGGAGTGT	1500						
Qy	1530	ATGTCACTTGGCTGGAGCACAACCTCAAGSGAAACTCCGAAGATTCTCT	1589						
Ds	1501	ATGTCACTTGGCTGGAGCACAACCTCAAGSGAAACTCCGAAGATTCTCT	1560						
Qy	1590	AGCATTTCTTCTTCTGATGCATATAGCACATTAATGGCACACACACAG	1649						
Ds	1561	AGCATTTCTTCTTCTGATGCATATAGCACATTAATGGCACACACACAG	1620						
Qy	1650	CAGATTACAGAGCAGCCCTTAGATCTCTTCTAGGGCACCCACAGAAA	1709						
Ds	1621	CAGATTACAGAGCAGCCCTTAGATCTCTTCTAGGGCACCCACAGAAA	1680						
Qy	1710	CTCTGGCCCAAGGGGAGGAGGACCTTTCAGCCCTTGGCTCACTCGAAT	1769						
Ds	1681	CTCTGGCCCAAGGGGAGGAGGACCTTTCAGCCCTTGGCTCACTCGAAT	1740						
Qy	1770	ATGAGGGTGGACCTTGTGTGGATTCTGTAGAGCCATGATGAGGTGAGG	1829						
Ds	1741	ATGAGGGTGGACCTTGTGTGGATTCTGTAGAGCCATGATGAGGTGAGG	1800						
Qy	1830	CCAGGGTGTGTGACAGCCCTTGGGTAAGTAAGTAACACTCTGAGGATT	1889						
Ds	1801	CCAGGGTGTGTGACAGCCCTTGGGTAAGTAAGTAACACTCTGAGGATT	1860						
Qy	1890	GAGATCTSAGAAACGCTGCTCTCATGTAGTAATCATCTATTCCCAAGG	1949						
Ds	1861	GAGATCTSAGAAACGCTGCTCTCATGTAGTAATCATCTATTCCCAAGG	1920						

QY	1950	TAGTCAAAAATAATTTGTTTATCCCCCCAAATTCCTATCTTTACAAATGGTGCTCATGAGAT	2009
DB	1921	TAGTCAAAAATAATTTGTTTATCCCCCCAAATTCCTATCTTTACAAATGGTGCTCATGAGAT	1980
QY	2010	TAGAACCCCTCTGTGTACTAATCAGCTTATCAACCAAGTCGAGACCTAGGAAGCTTAAT	2069
DB	1981	TAGAACCCCTCTGTGTACTAATCAGCTTATCAACCAAGTCGAGACCTAGGAAGCTTAAT	2040
QY	2070	GGATGGCAGACTGCTTAATTCGAGGAGGACTCAGAAGCAAAACCTACTTTCGGTTGCTT	2129
DB	2041	GGATGGCAGACTGCTTAATTCGAGGAGGACTCAGAAGCAAAACCTACTTTCGGTTGCTT	2100
QY	2130	TCATTATCTGCAACTTTAGAAAGAAATGATCTTTTTCCTCTGAAAGATACAAAGCT	2189
DB	2101	TCATTATCTGCAACTTTAGAAAGAAATGATCTTTTTCCTCTGAAAGATACAAAGCT	2160
QY	2190	CTGCAATTTGGTTGGAGTATTCCTACTCGACGCTGGAAGTTTAGCTTCACTGTGAATTT	2249
DB	2161	CTGCAATTTGGTTGGAGTATTCCTACTCGACGCTGGAAGTTTAGCTTCACTGTGAATTT	2220
QY	2250	AACAGAGAAAGTGCCTATAAGGGGGGGTTTAAAGAGACAATCCCATGATGCTGGGCCA	2309
DB	2221	AACAGAGAAAGTGCCTATAAGGGGGGGTTTAAAGAGACAATCCCATGATGCTGGGCCA	2280
QY	2310	ATGCTAAACAACAGGTCACGAACAACATAATGTTATAGAAGGAGCATCCCTCGACCATCTG	2369
DB	2281	ATGCTAAACAACAGGTCACGAACAACATAATGTTATAGAAGGAGCATCCCTCGACCATCTG	2340
QY	2370	ANTGAGAGTATGCTGTACCCCTTCCACCAAGAGTGGGGACACCTCTGCATATCTGCTCCC	2429
DB	2341	AATGAGAGTATGCTGTACCCCTTCCACCAAGAGTGGGGACACCTCTGCATATCTGCTCCC	2400
QY	2430	TCCTCTCTGTTAAGCCGCCAGGGAGCCCCATCCACCCAGTGGCTCTACACAGGGCAAT	2489
DB	2401	TCCTCTCTGTTAAGCCGCCAGGGAGCCCCATCCACCCAGTGGCTCTACACAGGGCAAT	2460
QY	2490	ACACACACCAAGATAGCTTCAGATCAACATGCATCAACATCAAGTGTTAATCTTCA	2549
DB	2461	ACACACACCAAGATAGCTTCAGATCAACATGCATCAACATCAAGTGTTAATCTTCA	2520
QY	2550	AGGTTTCTCTTCTTTTTCCTGTTTTTATTTGTTTTGCTTTTGCCTTTTCTTTTCTTT	2609
DB	2521	AGGTTTCTCTTCTTTTTCCTGTTTTTATTTGTTTTGCTTTTGCCTTTTCTTTTCTTT	2580
QY	2610	TTTTTGGTGGTGGGGCTACCAAACTGAGGCCCTAGAGCTAAAAATCATATAGAAATGA	2669
DB	2581	TTTTTGGTGGTGGGGCTACCAAACTGAGGCCCTAGAGCTAAAAATCATATAGAAATGA	2640
QY	2670	TGTTATCTGTGTGTGAGGAAGGCCAGCTGGCCTAAGTTCACACTTTTGTGCCAGTGS	2729
DB	2641	TGTTATCTGTGTGTGAGGAAGGCCAGCTGGCCTAAGTTCACACTTTTGTGCCAGTGS	2700
QY	2730	CCCTAGACTCCACCCAGCCAGCTCCCAAAATGAAAGACACCACTGTCAAGCAGCAGTCAAG	2789
DB	2701	CCCTAGACTCCACCCAGCCAGCTCCCAAAATGAAAGACACCACTGTCAAGCAGCAGTCAAG	2760
QY	2790	GAGTCTGATGTCAACCATCACTATTTTTTTCGATCATCTGTGCTTGCCTGCTCTCTTC	2849
DB	2761	GAGTCTGATGTCAACCATCACTATTTTTTTCGATCATCTGTGCTTGCCTGCTCTCTTC	2820
QY	2850	CACACCCGTGTGAGCTAATGCATTTGGGAAGCCAGGACAATGTTTGGCTGTTCTGCTTTGG	2909
DB	2821	CACACCCGTGTGAGCTAATGCATTTGGGAAGCCAGGACAATGTTTGGCTGTTCTGCTTTGG	2880
QY	2910	GTAAGGCACTCCCTGAAGCTCTGTGCTCTCCAGTATGGTGCTTTTCTTCTCTAACAG	2969
DB	2881	GTAAGGCACTCCCTGAAGCTCTGTGCTCTCCAGTATGGTGCTTTTCTTCTCTAACAG	2940
QY	2970	ATGCAATGTTTTCTTCAGATCAATAGTGATTCTTTAAATTAACCCAAAGACAGGCAT	3029
DB	2941	ATGCAATGTTTTCTTCAGATCAATAGTGATTCTTTAAATTAACCCAAAGACAGGCAT	3000
QY	3030	CCACAGTGTGTGAGCATGAATTCACAGGCTCGAATTTGTGTGAGTGTGAATAGTGAGATAAAA	3089

[illegible]

Search completed: October 24, 2003, 12:29:23
Job time : 3766 secs

XX	SP-061225-A1.
PK	
XX	C7-YAR-2001.
PD	
XX	30-AUG-1999;
PF	
XX	30-AUG-1999;
PF	
XX	30-AUG-1999;


```

DB 1681 TAGGCAACCCACAGAAACCCACAGACCCCGCTGGCCCCAGGCGGAGGAGGACACTTTTCAGC 1740
QY 1741 CTCCTGGCTCAGTCGAAATGTACAGAGCTTAGATGAGGGTGCACCTTTGGTTTGGATTCTGTA 1800
DB 1741 CTCCTGGCTCAGTCGAAATGTACAGAGCTTAGATGAGGGTGCACCTTTGGTTTGGATTCTGTA 1800
QY 1801 GAAGCCATGAGTGAAGTGGGAAGTGTGTTTCCAGGTTTGTGCCAGGCCCTGGTGAAGTAA 1860
DB 1801 GAAGCCATGAGTGAAGTGGGAAGTGTGTTTCCAGGTTTGTGCCAGGCCCTGGTGAAGTAA 1860
QY 1861 GACCTCTGAGGATTTCTCAGAGCCACACTTGAGATCTGAGGAGCGCTGCTCATGTAGTA 1920
DB 1861 GACCTCTGAGGATTTCTCAGAGCCACACTTGAGATCTGAGGAGCGCTGCTCATGTAGTA 1920
QY 1921 ATCATCTTATCCCAAGAGGCCCTCCGAGTAGTCAAAACTATTGTTTATCCGCCCAAAAT 1980
DB 1921 ATCATCTTATCCCAAGAGGCCCTCCGAGTAGTCAAAACTATTGTTTATCCGCCCAAAAT 1980
QY 1981 CCTATCTTTTAAATGTGCTGATGAGATTAACAACCCCTCTGTGTACTTAATCAGCTTATC 2040
DB 1981 CCTATCTTTTAAATGTGCTGATGAGATTAACAACCCCTCTGTGTACTTAATCAGCTTATC 2040
QY 2041 AACCAAGTGAGAACCTAGGAAAGCTAATTTGGATGCGACACTGCTTAATTCGAGGGAGGA 2100
DB 2041 AACCAAGTGAGAACCTAGGAAAGCTAATTTGGATGCGACACTGCTTAATTCGAGGGAGGA 2100
QY 2101 CTCAGAAGCCAAACCTACTCTCGTTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC 2160
DB 2101 CTCAGAAGCCAAACCTACTCTCGTTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC 2160
QY 2161 TTTTTTCCCTCGTAAAGATTAAGAAAGCTGCAATTTGGTTGGAGTATTCCTACTGCA 2220
DB 2161 TTTTTTCCCTCGTAAAGATTAAGAAAGCTGCAATTTGGTTGGAGTATTCCTACTGCA 2220
QY 2221 GCCTGGAAGTTAGCTTCACTGTGAATTTAAACAGAGAAAGTGCCTATTAAGGGGCGTTT 2280
DB 2221 GCCTGGAAGTTAGCTTCACTGTGAATTTAAACAGAGAAAGTGCCTATTAAGGGGCGTTT 2280
QY 2281 TTAAGAGCAATCCCATGATGTGCGCCCAATGCTTAACAACAGGGTCAAGAAACCAATGT 2340
DB 2281 TTAAGAGCAATCCCATGATGTGCGCCCAATGCTTAACAACAGGGTCAAGAAACCAATGT 2340
QY 2341 TTATAGAGGACATCCCTCGAGCATCTGAATGAGATGCTGACGCCCTTCCAGCACA 2400
DB 2341 TTATAGAGGACATCCCTCGAGCATCTGAATGAGATGCTGACGCCCTTCCAGCACA 2400
QY 2401 AGTGGGGACACCTGSCATATCTGCTCCCTCTCTGCTGTTAAGCCCGAGGAGCCCAT 2460
DB 2401 AGTGGGGACACCTGSCATATCTGCTCCCTCTCTGCTGTTAAGCCCGAGGAGCCCAT 2460
QY 2461 CCATCCAGTGGTCTTACAGACAGGGCAATACACACACAGATAGCCTTCAGATCAAC 2520
DB 2461 CCATCCAGTGGTCTTACAGACAGGGCAATACACACACAGATAGCCTTCAGATCAAC 2520
QY 2521 ATGCATCAGCTCAGCTGTTAATCTTCAAGGTTTCTTTTCTTTTCTCTGTTTCTTAT 2580
DB 2521 ATGCATCAGCTCAGCTGTTAATCTTCAAGGTTTCTTTTCTTTTCTCTGTTTCTTAT 2580
QY 2581 TGTATTGCTTTTGGTTTTTTTTTTTTTTTTTTTTTTTGGTGGTGGGGCTACCAAACTTGAG 2640
DB 2581 TGTATTGCTTTTGGTTTTTTTTTTTTTTTTTTTTTTTGGTGGTGGGGCTACCAAACTTGAG 2640
QY 2641 GCTAGAGCTAAATATCATAGAATGATGTTATCTTGTGTTGAGGAAAGGCCAGCT 2700
DB 2641 GCTAGAGCTAAATATCATAGAATGATGTTATCTTGTGTTGAGGAAAGGCCAGCT 2700
QY 2701 GGCTAAGTTTCACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
DB 2701 GGCTAAGTTTCACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTCCCAAAAT 2760
QY 2761 GAAAGACCCCTGTCTCAAGCAGCAGTCAAGAGTCTGATGTCACCCATCACTATTTTTT 2820
DB 2761 GAAAGACCCCTGTCTCAAGCAGCAGTCAAGAGTCTGATGTCACCCATCACTATTTTTT 2820

```

RESULT 2

```

AAAF5258
ID   AAAF5258 standard; cDNA; 3255 BP.
XX
AC   AAAF5258.
XX
DT   29-MAY-2001 (first entry)
XX
DE   Nucleotide sequence of a mutated murine parkin2 polypeptide.
XX
KW   Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW   Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW   multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW   brain tumour; head trauma; stroke; vascular irregularity;
KW   metabolic irregularity; ss.
XX
OS   Mus sp.
XX
KE   Key
XX   CDS
XX   /tag= a
XX   /product= "truncated parkin2"
XX
PE   PE1081225-A1.
XX
FC   07-MAR-2001.
XX
PF   30-AUG-1999; 99EP-018766.
XX
PR   30-AUG-1999; 99EP-018766.
XX
PA   (PICO-1) BIOFRONTIERA PHARM GMBH.
XX
PI   Liebert H;
XX
DR   WPI; 2001-212797/22.
XX   P-PSDB; AAB67511.
XX
PT   New polynucleotides encoding mouse parkin2 protein, useful for

```


CC the parkin2 gene cause Parkinson's disease in humans. The human
 CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
 CC and polynucleotides are useful for analysing neurodegenerative diseases.
 CC They are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis,
 CC Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease,
 CC and secondary causes including Parkinson's syndromes like toxins, drugs,
 CC brain tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.
 XX

Sequence 3255 BP; 853 A; 817 C; 809 G; 776 T; 0 other;
 Query Match 100.0%; Score 3253.4; DB 22; Length 3255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	CTCAGCGAGGGGAGGGGAGGAGGCTGTGATGACTAAACCTGACAGAAACGCTGTGGG	60
DB	1	CTCAGCGAGGGGAGGGGAGGAGGCTGTGATGACTAAACCTGACAGAAACGCTGTGGG	60
QY	61	AGGCTCGGGGGGGCGGAGTGCCTGGCTAGTCTCTGACCCGAGCAGCACCCGCG	120
DB	61	AGGCTCGGGGGGGCGGAGTGCCTGGCTAGTCTCTGACCCGAGCAGCACCCGCG	120
QY	121	CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGG	180
DB	121	CGGTGACCATGATAGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGG	180
QY	181	ATTCGACACGACATCTCCAGCTCAGAGAGTGGTGTCTGACGACAGAGGGTCCAG	240
DB	181	ATTCGACACGACATCTCCAGCTCAGAGAGTGGTGTCTGACGACAGAGGGTCCAG	240
QY	241	CTGACCAAGTCTGCTGTGATCTTTGCGGGGAGGAGTCTCCGAATCACTGACGCTTCAA	300
DB	241	CTGACCAAGTCTGCTGTGATCTTTGCGGGGAGGAGTCTCCGAATCACTGACGCTTCAA	300
QY	301	ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGAGACCCAGGAGAGTTC	360
DB	301	ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGAGACCCAGGAGAGTTC	360
QY	361	ATGAACACATATGATCTGGAGGGGACCAACCCAGAGCACCTCAGAGGCTCCATATGG	420
DB	361	ATGAACACATATGATCTGGAGGGGACCAACCCAGAGCACCTCAGAGGCTCCATATGG	420
QY	421	AGTCCAGAGGTTGACACGAGTGGACCTGAGCAGCCATACCTGCGCGTGGACTCTGTGG	480
DB	421	AGTCCAGAGGTTGACACGAGTGGACCTGAGCAGCCATACCTGCGCGTGGACTCTGTGG	480
QY	481	GGCTGGCGGTCACTCTGACACACAGATGAAGGATTCAGAGGACCCAGAGGCTCCAG	540
DB	481	GGCTGGCGGTCACTCTGACACACAGATGAAGGATTCAGAGGACCCAGAGGCTCCAG	540
QY	541	TTAAACCCACTACACAGCTTTTTCATCTACTGCAAGGCGCCCTGCCAAGAGTCCAGC	600
DB	541	TTAAACCCACTACACAGCTTTTTCATCTACTGCAAGGCGCCCTGCCAAGAGTCCAGC	600
QY	601	CTGCAAGAGCTCCGAGTTCAGTGTGGACCTGCAACAGCAACCTCAGCTTGGCCAGG	660
DB	601	CTGCAAGAGCTCCGAGTTCAGTGTGGACCTGCAACAGCAACCTCAGCTTGGCCAGG	660
QY	661	GCCCATCTTGTGGGAGGATGTCTTAATTCGAACCGGATGAGTGGTGGAGTCTC	720
DB	661	GCCCATCTTGTGGGAGGATGTCTTAATTCGAACCGGATGAGTGGTGGAGTCTC	720
QY	721	CAGACTGCCCTGGAAACAGAGCTGAATTTCTTTAAATGTGGAGCACACCCAACTCAG	780
DB	721	CAGACTGCCCTGGAAACAGAGCTGAATTTCTTTAAATGTGGAGCACACCCAACTCAG	780
QY	781	ACAAGGACAGCTGGTAGCTTGAACCTGATACCAAGCAACAGCGCAGCATCCCTTGA	840
DB	781	ACAAGGACAGCTGGTAGCTTGAACCTGATACCAAGCAACAGCGCAGCATCCCTTGA	840

QY	841	TAGGTCACAGATGTACAGAGCCCTGTCTCTGGTCTCCAGTGTACACACCGTCACTGA	900
DB	841	TAGGTCACAGATGTACAGAGCCCTGTCTCTGGTCTCCAGTGTACACACCGTCACTGA	900
QY	901	TCTGTTGGAGTCTTTTCCACTTGTATTGTGTACAGAGTCAACGATCGGAGTGTGTC	960
DB	901	TCTGTTGGAGTCTTTTCCACTTGTATTGTGTACAGAGTCAACGATCGGAGTGTGTC	960
QY	961	ACGATGCTCAACTTGGCTACTCTCTGCTGCTGTAGCTGGCTGTCCCACTCTCTGATTA	1020
DB	961	ACGATGCTCAACTTGGCTACTCTCTGCTGCTGTAGCTGGCTGTCCCACTCTCTGATTA	1020
QY	1021	AAGAGCTCCATCACTTCAAGATCCITGGAGAGAGAGTACACTAGGTACACAGCAGTATG	1080
DB	1021	AAGAGCTCCATCACTTCAAGATCCITGGAGAGAGAGTACACTAGGTACACAGCAGTATG	1080
QY	1081	GGCCCGAGGAGTCTGCTGCTGCAAAATGGGAGTGTGCTGTGCTGCTGCTGCTGCTG	1140
DB	1081	GGCCCGAGGAGTCTGCTGCTGCAAAATGGGAGTGTGCTGTGCTGCTGCTGCTGCTG	1140
QY	1141	CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
DB	1141	CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
QY	1201	GCTGCGGCTTGTGTTTCTGCGGGAGTGTAAAGAGCAATACCATGAAGGGGATTCGACT	1260
DB	1201	GCTGCGGCTTGTGTTTCTGCGGGAGTGTAAAGAGCAATACCATGAAGGGGATTCGACT	1260
QY	1261	GCTGCTGCGAGCCTCAGAGGACCTCTCTCAGGCTTACAGGCTGAGAGGAGGAGGCTG	1320
DB	1261	GCTGCTGCGAGCCTCAGAGGACCTCTCTCAGGCTTACAGGCTGAGAGGAGGAGGCTG	1320
QY	1321	AGGAAGCTCTGCTGGAGAGGAGGCTCTCAAGAGAAATATCAAGAGAGCAACCAAGCTTGC	1380
DB	1321	AGGAAGCTCTGCTGGAGAGGAGGCTCTCAAGAGAAATATCAAGAGAGCAACCAAGCTTGC	1380
QY	1381	CTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1440
DB	1381	CTGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1440
QY	1441	CCAGTCCAGGCTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
DB	1441	CCAGTCCAGGCTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
QY	1501	GAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560
DB	1501	GAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560
QY	1561	GGGAAGCTCCGAGAGTTCCTACCTTCTTAGGCAATCTCTTCTGATGATGATGATGATG	1620
DB	1561	GGGAAGCTCCGAGAGTTCCTACCTTCTTAGGCAATCTCTTCTGATGATGATGATGATG	1620
QY	1621	CATAAATGGGCAACACACAGAGCTGCAATATACAGAGCAGCCCTAGATCTTCTT	1680
DB	1621	CATAAATGGGCAACACACAGAGCTGCAATATACAGAGCAGCCCTAGATCTTCTT	1680
QY	1681	TAGGCGCCACAGAGAAACACAGAGCAGCTGGCCCGGAGGGGAGGAGGAGGAGGAGG	1740
DB	1681	TAGGCGCCACAGAGAAACACAGAGCAGCTGGCCCGGAGGGGAGGAGGAGGAGGAGG	1740
QY	1741	CTCTGGCTCACTCGAATGTGACAGCTTAGATGAGGTGCACTTGGTTGGTATCTGTA	1800
DB	1741	CTCTGGCTCACTCGAATGTGACAGCTTAGATGAGGTGCACTTGGTTGGTATCTGTA	1800
QY	1801	GAGGCTAGTGGAGTGGAGAGTGTGTTTCCAGGCTTGTGTCAGCCCTGGTGAAGTAA	1860
DB	1801	GAGGCTAGTGGAGTGGAGAGTGTGTTTCCAGGCTTGTGTCAGCCCTGGTGAAGTAA	1860
QY	1861	CAGCTCTGAGGATCTCAGAGAGCACTTGAATGTGAGAGAGGCTGCTCTCATGTAGTA	1920
DB	1861	CAGCTCTGAGGATCTCAGAGAGCACTTGAATGTGAGAGAGGCTGCTCTCATGTAGTA	1920

QY	1921	ATCATCTATTCCAAAGGGCCCCCTCAGTAGTCAAAACTATTTGTTTATTCGCCCCCAAT	1980
DB	1921	ATCATCTATTCCAAAGGGCCCCCTCAGTAGTCAAAACTATTTGTTTATTCGCCCCCAAT	1980
QY	1981	CGTATCTTTTACAATGGTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC	2040
DB	1981	CGTATCTTTTACAATGGTGTGATGAGATTACAAACCCCTCTGTGTACTAATCAGCTTATC	2040
QY	2041	AACCAATGAGAACCTAGGAAAGCTAATTTGGATGGCAGACTGCTTAAATCGCAGGGAGGA	2100
DB	2041	AACCAATGAGAACCTAGGAAAGCTAATTTGGATGGCAGACTGCTTAAATCGCAGGGAGGA	2100
QY	2101	CTCAGAACCAACCTACTTCCGTTGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2160
DB	2101	CTCAGAACCAACCTACTTCCGTTGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2160
QY	2161	TTTTTTCGCCCTGAAAAGATAACAAAGTCTGCAATTTGGTTTGGAGTATTCCTACTGCA	2220
DB	2161	TTTTTTCGCCCTGAAAAGATAACAAAGTCTGCAATTTGGTTTGGAGTATTCCTACTGCA	2220
QY	2221	GCCTGGAGTTTAGCTTCTACTGTGSAATTTTAAACAGAGAAAGTGCCTATTAAGGGGGCGTT	2280
DB	2221	GCCTGGAGTTTAGCTTCTACTGTGSAATTTTAAACAGAGAAAGTGCCTATTAAGGGGGCGTT	2280
QY	2281	TTAAGAGACAATCCCATGATGCTGGCGCAATGTATAACACAGGGTCAAGAAACACAATGT	2340
DB	2281	TTAAGAGACAATCCCATGATGCTGGCGCAATGTATAACACAGGGTCAAGAAACACAATGT	2340
QY	2341	TTATAGAGAGAGATCCCTCGACATCTGTAATGAGAGTATGCTCTGACCCCTTCCACACA	2400
DB	2341	TTATAGAGAGAGATCCCTCGACATCTGTAATGAGAGTATGCTCTGACCCCTTCCACACA	2400
QY	2401	AGTGGGACACCTCTGCATATCTGCTCCCTCTCTGCTGTATGAGCCCCAGGAGAGCCCAT	2460
DB	2401	AGTGGGACACCTCTGCATATCTGCTCCCTCTCTGCTGTATGAGCCCCAGGAGAGCCCAT	2460
QY	2461	CCACCCAGTGGTCTTACAGACAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC	2520
DB	2461	CCACCCAGTGGTCTTACAGACAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC	2520
QY	2521	ATGCATCAGACTCAAGTGTAACTTTTCAAGGTTTCTCTTCTTTTCTGTTTTTTATT	2580
DB	2521	ATGCATCAGACTCAAGTGTAACTTTTCAAGGTTTCTCTTCTTTTCTGTTTTTTATT	2580
QY	2581	TGTTTCTGTTTTGCTTTTTTTTTTTTTTTTTTTTGGTGGTGGGCTACCAAACTGAG	2640
DB	2581	TGTTTCTGTTTTGCTTTTTTTTTTTTTTTTTTTTGGTGGTGGGCTACCAAACTGAG	2640
QY	2641	SCCTAGAGCTTAAATATCATATAGAAATGATGTTATCTGTGGTGTGAGGAAAGGCCAGCT	2700
DB	2641	SCCTAGAGCTTAAATATCATATAGAAATGATGTTATCTGTGGTGTGAGGAAAGGCCAGCT	2700
QY	2701	GGCTAAGTTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTGCCAAAT	2760
DB	2701	GGCTAAGTTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCCAGCCAGCTGCCAAAT	2760
QY	2761	GAAGAAGCAGCTGTGAAGCAGCAGTCAAGAGTCTGATGTCAACCCATCACTATTTTTTTT	2820
DB	2761	GAAGAAGCAGCTGTCAAGCAGCAGTCAAGAGTCTGATGTCAACCCATCACTATTTTTTTT	2820
QY	2821	CCATCATGTGCTTGGCTCTGCCCTTCCACACCCGCTGTGACGTATCCATTTGGGAAG	2880
DB	2821	CCATCATGTGCTTGGCTCTGCCCTTCCACACCCGCTGTGACGTATCCATTTGGGAAG	2880
QY	2881	CCAGGACAAATGTTTGTCTGTTTTGGGTAAAGGAGTCCCTGAAAGCTCTGTGGCTCT	2940
DB	2881	CCAGGACAAATGTTTGTCTGTTTTGGGTAAAGGAGTCCCTGAAAGCTCTGTGGCTCT	2940
QY	2941	CCAGTATGTCCTTTTCCCTTCCTTAAAGATGCAATGTTTCTTTCAGAAATACATAGT	3000
DB	2941	CCAGTATGTCCTTTTCCCTTAAAGATGCAATGTTTCTTTCAGAAATACAAATAGT	3000
QY	3001	ATTCTTAAATTAACCCAAAAGACAGGATCCACAGTGTGTGAGCATGAATCAGACCTGCG	3060

[illegible]

CC and secondary causes inducing Parkinson's syndromes like toxins, drugs,
 CC brain tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 XX protein.

SQ Sequence 3255 BP; 853 A; 818 C; 808 G; 776 T; 0 other;

Query Match	100.0%	Score 3253.4	DB 22	Length 3255;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 3254;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy 1	CTCAGCGAGGGAAGGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG	60		
Db 1	CTCAGCGAGGGAAGGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG	60		
Qy 61	AGGCTCGGGCGGGCGCCAGTCCCGCGGTAGTCTCTTCGACCGGACGACACCGCCGCG	120		
Db 61	AGGCTCGGGCGGGCGCCAGTCCCGCGGTAGTCTCTTCGACCGGACGACACCGCCGCG	120		
Qy 121	CGGTGACCATGATAGTGTCTTCAGGTTCAACTCCAGCTATGGCTTCCCACTGGAGGTGG	180		
Db 121	CGGTGACCATGATAGTGTCTTCAGGTTCAACTCCAGCTATGGCTTCCCACTGGAGGTGG	180		
Qy 181	ATTCTGACACGAGCATCTTGAGCTCAAGAGTGGTGTCTTAAGCGGACGAGGGGTTCAG	240		
Db 181	ATTCTGACACGAGCATCTTGAGCTCAAGAGTGGTGTCTTAAGCGGACGAGGGGTTCAG	240		
Qy 241	CTGACACGCTGGTGTATTTTCCGGGAGGAGCTTCCGATCACTGACGGTTCMAA	300		
Db 241	CTGACACGCTGGTGTATTTTCCGGGAGGAGCTTCCGATCACTGACGGTTCMAA	300		
Qy 301	ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTTC	360		
Db 301	ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTTC	360		
Qy 361	ATGAACAAATGCACTCTGGAGGGGACGACCCGAGAGACCTCAGAGGGTCCATATGGG	420		
Db 361	ATGAACAAATGCACTCTGGAGGGGACGACCCGAGAGACCTCAGAGGGTCCATATGGG	420		
Qy 421	AGTCCAGGAGCTTGACAGAGTGGACCTTGACAGCCATACCTCCCGGTGGATCTGTGG	480		
Db 421	AGTCCAGGAGCTTGACAGAGTGGACCTTGACAGCCATACCTCCCGGTGGATCTGTGG	480		
Qy 481	GGCTGCGGTGATCTTGACACACAGCATTAAGAGGATTCAGAGGAGCCAGAGGTCCAG	540		
Db 481	GGCTGCGGTGATCTTGACACACAGCATTAAGAGGATTCAGAGGAGCCAGAGGTCCAG	540		
Qy 541	TTAAACCCACCTTACACAGCTTTTTCATCTACTGCAAGGCCCTGCCACAAGGTCCAGC	600		
Db 541	TTAAACCCACCTTACACAGCTTTTTCATCTACTGCAAGGCCCTGCCACAAGGTCCAGC	600		
Qy 601	CTGGAAGCTCGAGTTCAGTGTGGCACCTGCAACAGAGAACCTCAGCTTGGCCGAGG	660		
Db 601	CTGGAAGCTCGAGTTCAGTGTGGCACCTGCAACAGAGAACCTCAGCTTGGCCGAGG	660		
Qy 661	GCCATCTTGTGGGACATGCTTAAATCCAAACCGGATGAGTGGTGGCCAGTCTC	720		
Db 661	GCCATCTTGTGGGACATGCTTAAATCCAAACCGGATGAGTGGTGGCCAGTCTC	720		
Qy 721	CAGACTGCGCTGGAAACGAGCTGAATTTTCTTTTAAATGTGAGGACACCCCAACCTCAG	780		
Db 721	CAGACTGCGCTGGAAACGAGCTGAATTTTCTTTTAAATGTGAGGACACCCCAACCTCAG	780		
Qy 781	ACAGGACAGCTGCTAGCTTGAACCTGATCACAGCAAGCGGCGAGGATCCCTTGA	840		
Db 781	ACAGGACAGCTGCTAGCTTGAACCTGATCACAGCAAGCGGCGAGGATCCCTTGA	840		
Qy 841	TAGCGTGCACAGATGTCAGGAGCCCTGCTCGGTCTCCAGTGTAAACACCGTCCAGCTGA	900		
Db 841	TAGCGTGCACAGATGTCAGGAGCCCTGCTCGGTCTCCAGTGTAAACACCGTCCAGCTGA	900		
Qy 901	TCTGTTGAGCTGTTTCCACTGTATTGTGTCAAGACTCAACGATGGGAGGTTCCTCC	960		
Db 901	TCTGTTGAGCTGTTTCCACTGTATTGTGTCAAGACTCAACGATGGGAGGTTCCTCC	960		

Qy 901	TCTGTTTGGACTGTTTCCACTGTATTGTGTCAAGACTCAACGATCGCAGTTTGTCC	960	Db
Qy 961	ACGATGCTCAACTTGGCTACTCCTCCCGTGTACTGCTGTCCCAACTCCCTGATTA	1020	Qy
Db 961	ACGATGCTCAACTTGGCTACTCCTCCCGTGTACTGCTGTCCCAACTCCCTGATTA	1020	Db
Qy 1021	AAGAGCTCCATCACTTCCAGGATCCITGGAGAGAGCAGTACACTAGGTACCAAGCATG	1080	Qy
Db 1021	AAGAGCTCCATCACTTCCAGGATCCITGGAGAGAGCAGTACACTAGGTACCAAGCATG	1080	Db
Qy 1081	GGGCGGAGGAATGCGTCTCCAAATGAGAGGTGTGTGTGCCCCCTCTGCTGTGGAG	1140	Qy
Db 1081	GGGCGGAGGAATGCGTCTCCAAATGAGAGGTGTGTGTGCCCCCTCTGCTGTGGAG	1140	Db
Qy 1141	CTGCACTGCTCCGTGACAGGGCCAGGGAAGTCACTGGGAAAGGGGCAAGGGCTCG	1200	Qy
Db 1141	CTGCACTGCTCCGTGACAGGGCCAGGGAAGTCACTGGGAAAGGGGCAAGGGCTCG	1200	Db
Qy 1201	CTGCACTGCTCCGTGACAGGGCCAGGGAAGTCACTGGGAAAGGGGCAAGGGCTCG	1260	Qy
Db 1201	CTGCACTGCTCCGTGACAGGGCCAGGGAAGTCACTGGGAAAGGGGCAAGGGCTCG	1260	Db
Qy 1261	CAGTGTGGAACCTCTAGGAGCCACTTCTCAGCCCTACAGGCTGGACAAAGAGCCGTG	1320	Qy
Db 1261	CAGTGTGGAACCTCTAGGAGCCACTTCTCAGCCCTACAGGCTGGACAAAGAGCCGTG	1320	Db
Qy 1321	AGCAAGCTCGCTGGAGGAGGAGTCAAGGAAACCATCAAGAAAGCAACCAAGCCCTGTG	1380	Qy
Db 1321	AGCAAGCTCGCTGGAGGAGGAGTCAAGGAAACCATCAAGAAAGCAACCAAGCCCTGTG	1380	Db
Qy 1381	CTGCTGCAACGTCGCTGGAATTTGAAAGAAACCGAGGATGATGCAACATGAAGTCTCAGC	1440	Qy
Db 1381	CTGCTGCAACGTCGCTGGAATTTGAAAGAAACCGAGGATGATGCAACATGAAGTCTCAGC	1440	Db
Qy 1441	CCGAGTCAAGCTGGAGTGGTGTGGAATGTGCTGTGGAACCGGAGCTGATGG	1500	Qy
Db 1441	CCGAGTCAAGCTGGAGTGGTGTGGAATGTGCTGTGGAACCGGAGCTGATGG	1500	Db
Qy 1501	GAGATCACTGGTGTGAGAGAGATGTCACTTGGCCCTTGGAGCGGACGACACCTCAA	1560	Qy
Db 1501	GAGATCACTGGTGTGAGAGAGATGTCACTTGGCCCTTGGAGCGGACGACACCTCAA	1560	Db
Qy 1561	GGGAAATCTCGAAGATCTTACCTTCTTACGATTCCTTCTTCTGATSCATAGCA	1620	Qy
Db 1561	GGGAAATCTCGAAGATCTTACCTTCTTACGATTCCTTCTTCTGATSCATAGCA	1620	Db
Qy 1621	CATAAATGGGCAACACACAGGCTGAGATTCAGAGAGCGGCCCTAGATTCCTTC	1680	Qy
Db 1621	CATAAATGGGCAACACACAGGCTGAGATTCAGAGAGCGGCCCTAGATTCCTTC	1680	Db
Qy 1681	TAGGGCAACCCACAGAAACACAGCACCCTGGGCCAGGGGAGGAGGACCTTCAGC	1740	Qy
Db 1681	TAGGGCAACCCACAGAAACACAGCACCCTGGGCCAGGGGAGGAGGACCTTCAGC	1740	Db
Qy 1741	CTCTGGCTCACTCGAATGTCTCAGAGCTTAGATGAGGCTGCACTTTGGTTTGGATTCGTA	1800	Qy
Db 1741	CTCTGGCTCACTCGAATGTCTCAGAGCTTAGATGAGGCTGCACTTTGGTTTGGATTCGTA	1800	Db
Qy 1801	GAGCCATGATGAGGTGGGAAAGTGTTCAGAGGTGTTCGACAGCCCTGGGTAGATTA	1860	Qy
Db 1801	GAGCCATGATGAGGTGGGAAAGTGTTCAGAGGTGTTCGACAGCCCTGGGTAGATTA	1860	Db
Qy 1861	CACCTCTGAGGATCTCAGAAAGCAGCTTGGAGAGCGCTGTCTCATGTAGTA	1920	Qy
Db 1861	CACCTCTGAGGATCTCAGAAAGCAGCTTGGAGAGCGCTGTCTCATGTAGTA	1920	Db
Qy 1921	ATCATCTATTCGAAAGGCGCCCTGAGTGTAGTCAAAACCTATTGTTTATCCGCCCAAT	1980	Qy
Db 1921	ATCATCTATTCGAAAGGCGCCCTGAGTGTAGTCAAAACCTATTGTTTATCCGCCCAAT	1980	Db
Qy 1981	CCTATCTTTCAAAATGCTGTGATGAGATTAACAGCCCTCTGTGTACTATCATGCTATC	2040	Qy
Db 1981	CCTATCTTTCAAAATGCTGTGATGAGATTAACAGCCCTCTGTGTACTATCATGCTATC	2040	Db

Query Match									
Best Local Similarity 99.6%; Score 3243; DB 22; Length 3254;									
Matches 3254; Conservative 0; Mismatches 0; Indels 1; Gaps 1;									
QY	1	CTCAGCGAGGGAAGGGGGAGGAGGCGCTGGATGACTAAACCTGCAGAAACCGTGGTGGG	60						
DB	1	CTCAGCGAGGGAAGGGGGAGGAGGCGCTGGATGACTAAACCTGCAGAAACCGTGGTGGG	60						
QY	61	AGGCTCGGGGGGGCCAGTGGCCGGCTAGTCTCTTCGACCGGAGCCACACCGCC	120						
DB	61	AGGCTCGGGGGGGCCAGTGGCCGGCTAGTCTCTTCGACCGGAGCCACACCGCC	120						
QY	121	CGGTGACCATGATAGTGTGTCAGGTTCACACTCCAGCTATGGCTTCCCACTGAGGTG	180						
DB	121	CGGTGACCATGATAGTGTGTCAGGTTCACACTCCAGCTATGGCTTCCCACTGAGGTG	180						
QY	181	ATCTGACACAGCATCTTGAGGCTCAAGGAAGTGGTTCGTAACGACAGGGGTCCAG	240						
DB	181	ATCTGACACAGCATCTTGAGGCTCAAGGAAGTGGTTCGTAACGACAGGGGTCCAG	240						
QY	241	CTGACCGACTCGGTGTGATTTTGGCGGGAAGAGCTTCCGATCACTGACCGTTCAA	300						
DB	241	CTGACCGACTCGGTGTGATTTTGGCGGGAAGAGCTTCCGATCACTGACCGTTCAA	300						
QY	361	ACTGTGACCTGGAACAACAGAGTATTTGACACATAGTACAGAGACACGAGGAGAGTC	360						
DB	361	ACTGTGACCTGGAACAACAGAGTATTTGACACATAGTACAGAGACACGAGGAGAGTC	360						
QY	421	ATGAAACAATGCACTCTGGAGGGGAGCAACCCGAGAGACCTCAGAGGGTCCATATGGG	420						
DB	421	ATGAAACAATGCACTCTGGAGGGGAGCAACCCGAGAGACCTCAGAGGGTCCATATGGG	420						
QY	481	GGCTGGCGGTCACTCTGACACAGACAGTAAAGAGGATTCAGAGGAGCCAGAGGTCCAG	540						
DB	481	GGCTGGCGGTCACTCTGACACAGACAGTAAAGAGGATTCAGAGGAGCCAGAGGTCCAG	540						
QY	541	TTAAACCCACCTACACAGCTTTTTCATCTACTCGAAGGCCCTGCCAAGGTCCAGC	600						
DB	541	TTAAACCCACCTACACAGCTTTTTCATCTACTCGAAGGCCCTGCCAAGGTCCAGC	600						
QY	601	CTGGAAGGCTCGAGTTCAGTGTGACCTGCAACCAAGCAACCCCTCACCTTGGCCCGAGG	660						
DB	601	CTGGAAGGCTCGAGTTCAGTGTGACCTGCAACCAAGCAACCCCTCACCTTGGCCCGAGG	660						
QY	661	GCCGATCTGCTGGAGCATGCTTAATTCGAAACCGGATGAGTGGTGGAGTCCAGTCTC	720						
DB	661	GCCGATCTGCTGGAGCATGCTTAATTCGAAACCGGATGAGTGGTGGAGTCCAGTCTC	720						
QY	721	CAGACTGCCCTGGAAACACAGAGCTGAATTTTCTTTAAATGAGGACACACCCACCTCAG	780						
DB	721	CAGACTGCCCTGGAAACACAGAGCTGAATTTTCTTTAAATGAGGACACACCCACCTCAG	780						
QY	781	ACAAGGACACGTCGGTAGCTTTGAACCTGATCACAGCAACAGCGCGAGCATCCCTTGCA	840						
DB	781	ACAAGGACACGTCGGTAGCTTTGAACCTGATCACAGCAACAGCGCGAGCATCCCTTGCA	840						
QY	841	TAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCGAGTGAACCCACCGTCACGTGA	900						
DB	841	TAGCGTGCACAGATGTCAGGAGCCCTGTCCTGGTCTTCGAGTGAACCCACCGTCACGTGA	900						
QY	901	CTGTGTTGACATGTTTCCACTTGTATGTGTCACAGAGCTCAACGATCGGAGGTGTGTC	960						
DB	901	CTGTGTTGACATGTTTCCACTTGTATGTGTCACAGAGCTCAACGATCGGAGGTGTGTC	960						
QY	961	AGGATGCTCAACTTGGCTACTTCCCTGCGGTGTGTAGCTGGCTGTCCCAACTCCCTGATTA	1020						
DB	961	AGGATGCTCAACTTGGCTACTTCCCTGCGGTGTGTAGCTGGCTGTCCCAACTCCCTGATTA	1020						

QY	1021	AAGAGCTCCATCACTTCCAGGATCCTTGGAGAAAGCAGTACACTAGGTACGAGCAGTATG	1080						
DB	1021	AAGAGCTCCATCACTTCCAGGATCCTTGGAGAAAGCAGTACACTAGGTACGAGCAGTATG	1080						
QY	1081	GGCCCGAGGAATGGTGTGCAAAATGGAGAGTGTCTGTGCCCCCGTCTGGTGTGTGAG	1140						
DB	1081	GGCCCGAGGAATGGTGTGCAAAATGGAGAGTGTCTGTGCCCCCGTCTGGTGTGTGAG	1140						
QY	1141	CTGGACTCTACCTGCAACAGGGCCAGAGGAAGTCACTGCGAAAGGGGGAACGGCTGG	1200						
DB	1141	CTGGACTCTACCTGCAACAGGGCCAGAGGAAGTCACTGCGAAAGGGGGAACGGCTGG	1200						
QY	1201	CTGCGGGTGTGTTTCTGCGGGGACTGTAAGSAGGATACCATGAAAGGGGATTCGACT	1260						
DB	1201	CTGCGGGTGTGTTTCTGCGGGGACTGTAAGSAGGATACCATGAAAGGGGATTCGACT	1260						
QY	1261	CACCTGCTCGAACCTCAGAGGACCTTCTGAGGCTTACAGGCTGGAGCAAGAGCGCTG	1320						
DB	1261	CACCTGCTCGAACCTCAGAGGACCTTCTGAGGCTTACAGGCTGGAGCAAGAGCGCTG	1320						
QY	1321	AGCAAGCTCGTGGAGGAGCGCTCCAGAGGAACCATCAAGAAACACCAAGCTTGTG	1380						
DB	1321	AGCAAGCTCGTGGAGGAGCGCTCCAGAGGAACCATCAAGAAACACCAAGCTTGTG	1380						
QY	1381	CTGCGTCAACCTGCAATTTGAAAAAAGGAGGATGTATGCATATGAAATGTCTCAGC	1440						
DB	1381	CTGCGTCAACCTGCAATTTGAAAAAAGGAGGATGTATGCATATGAAATGTCTCAGC	1440						
QY	1441	TCCAGTCAAGCTGAGTGTGCTGCGGACTGTGCTGAGTGGAGCGGAGCTTGTG	1500						
DB	1441	TCCAGTCAAGCTGAGTGTGCTGCGGACTGTGCTGAGTGGAGCGGAGCTTGTG	1500						
QY	1501	SAGATCACTGCTTTCAGCTGTAGAGAGAGATTTACCTTGGCCCTGGAGCAACCTCAA	1560						
DB	1501	SAGATCACTGCTTTCAGCTGTAGAGAGAGATTTACCTTGGCCCTGGAGCAACCTCAA	1560						
QY	1561	GGGAAACTCCGAGATTCCTACCTCCTTACGCAATTTCTTCTCTGATGCATATAGCA	1620						
DB	1561	GGGAAACTCCGAGATTCCTACCTCCTTACGCAATTTCTTCTCTGATGCATATAGCA	1620						
QY	1621	CATAAATGCGCACACACAAACACAGGCTGCGAGATTACAGAGAGCGCCCTAGATCCCTTC	1680						
DB	1621	CATAAATGCGCACACACAAACACAGGCTGCGAGATTACAGAGAGCGCCCTAGATCCCTTC	1680						
QY	1681	TAGGCAACCCACAGAAAAACACAGACACCCCGTGGCCAGGGGGAGGAGGACCTTCAGC	1740						
DB	1681	TAGGCAACCCACAGAAAAACACAGACACCCCGTGGCCAGGGGGAGGAGGACCTTCAGC	1740						
QY	1741	CTCTGGCTCACTCGAATGTGAGAGCTTAGATGAGGCTGCACCTTGGTTTGGATTCGTA	1800						
DB	1741	CTCTGGCTCACTCGAATGTGAGAGCTTAGATGAGGCTGCACCTTGGTTTGGATTCGTA	1800						
QY	1801	GAAGCCATGAGTGGTGGGAAGTCTTCCAGGCTTGTTCGACGCGCTGGTAAAGTAA	1860						
DB	1801	GAAGCCATGAGTGGTGGGAAGTCTTCCAGGCTTGTTCGACGCGCTGGTAAAGTAA	1860						
QY	1861	CACCTCTGAGGATCTCAGAGACACACTTGAGATCTGAGSAGAGCTGCTCTGATGATA	1920						
DB	1861	CACCTCTGAGGATCTCAGAGACACACTTGAGATCTGAGSAGAGCTGCTCTGATGATA	1920						
QY	1921	ATCATCTATTCGCAAGGCGCCCGCTGGAGTAGTGAAAAACTATTTGTTATCGCCGCAAT	1980						
DB	1921	ATCATCTATTCGCAAGGCGCCCGCTGGAGTAGTGAAAAACTATTTGTTATCGCCGCAAT	1980						
QY	1981	CTATCTCTTACAAATGGTGTGATGAGATTCAGATTCGCTCTGTGATCTAATCAGCTTATC	2040						
DB	1981	CTATCTCTTACAAATGGTGTGATGAGATTCAGATTCGCTCTGTGATCTAATCAGCTTATC	2040						
QY	2041	AGCAAGTGAAGAACTAGAAAGCTTAATGATGGAGACACTGCTTAAATGGAGGAGCA	2100						
DB	2041	AGCAAGTGAAGAACTAGAAAGCTTAATGATGGAGACACTGCTTAAATGGAGGAGCA	2100						
QY	2101	CTCAGAAAGCAACCTACTTCTGCTGCTTCTCATTATCTGCAACTTTAGAAAGAAATGATC	2160						

DB	2100	CTCAGAGGCGAAACCCACITTCGCGTGGTTCATATCTGCAATCTTAGAAGAAGATGATC	2159
QY	2161	TTTTTTTCCCCCTGAAAAGATAACAAAGCTGCGAATTTGGTTTGGAGTATTCCTACTGCA	2220
DB	2160	TTTTTTTCCCCCTGAAAAGATAACAAAGCTGCGAATTTGGTTTGGAGTATTCCTACTGCA	2219
QY	2221	GCCTGGAAGTTTACCTTCACCTGTGAATTTAAACAGAGAAAGTGCTATAAAGGGGGCGTTT	2280
DB	2220	GCCTGGAAGTTTACCTTCACCTGTGAATTTAAACAGAGAAAGTGCTATAAAGGGGGCGTTT	2279
QY	2281	TTAAGAGACAATCCCATGATGCTGGCGCAATGCTAACACAGGGTCAAGAGAAACACAATGT	2340
DB	2280	TTAAGAGACAATCCCATGATGCTGGCGCAATGCTAACACAGGGTCAAGAGAAACACAATGT	2339
QY	2341	TTTATGAAGGAGGATCCCTCGACCATCTGAATCAGAGTATGCTGTGACCCCTTCCACACCA	2400
DB	2340	TTTATGAAGGAGGATCCCTCGACCATCTGAATCAGAGTATGCTGTGACCCCTTCCACACCA	2399
QY	2401	AGTGGGACACTCTGCATATCTGCTCCCTCTCTGCTGTATAGCCGCCAGGGAGGCCCAT	2460
DB	2400	AGTGGGACACTCTGCATATCTGCTCCCTCTCTGCTGTATAGCCGCCAGGGAGGCCCAT	2459
QY	2461	CCACCCAGTGGTCTTACAGACAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC	2520
DB	2460	CCACCCAGTGGTCTTACAGACAGGGCAATACACACACCAAGATAGCCTTCAGATCAAC	2519
QY	2521	ATGCATCAGACTCAAGTGTTAACTTTTCAGGTTTCTTTTCTTTTCTTTTCTTTTATTT	2580
DB	2520	ATGCAACACACTCAAGTGTTAACTTTTCAGGTTTCTTTTCTTTTCTTTTCTTTTATTT	2579
QY	2581	TGTTTTTGCTTTTGCTTTTTTTTTTTTTTTTTTTTGGTGGTGGGCTACCAAACTTGAG	2640
DB	2580	TGTTTTTGCTTTTGCTTTTTTTTTTTTTTTTTTTTGGTGGTGGGCTACCAAACTTGAG	2639
QY	2641	GCCTAGAGCTAAAAATCATAAGAAATGATGTATCTTTGGTGTGAGGAAAGGCGAGCT	2700
DB	2640	GCCTAGAGCTAAAAATCATAAGAAATGATGTATCTTTGGTGTGAGGAAAGGCGAGCT	2699
QY	2701	GGCCTAAGTTCACACTTTTGCTCCAGTGGCGCTAGACTCCACCCAGCGCAGCTCCCAAAAT	2760
DB	2700	GGCCTAAGTTCACACTTTTGCTCCAGTGGCGCTAGACTCCACCCAGCGCAGCTCCCAAAAT	2759
QY	2761	GAAAGACCACTGTGCAAGCAGGAGTCAGAGCTGATGTGACCCCATCAGTATTTTTTTT	2820
DB	2760	GAAAGACCACTGTGCAAGCAGGAGTCAGAGCTGATGTGACCCCATCAGTATTTTTTTT	2819
QY	2821	CCATCATTTGGTTCGCTCTGCTCTTCCACACCCGCTGTGAGGTAAAGCATGTGGCAAG	2880
DB	2820	CCATCATTTGGTTCGCTCTGCTCTTCCACACCCGCTGTGAGGTAAAGCATGTGGCAAG	2879
QY	2881	CCAGGACAAATGTTTGGCTTCTGCTTTGGSTAAAGGACCTCCCTGAAGCTCTGTGCTCT	2940
DB	2880	CCAGGACAAATGTTTGGCTTCTGCTTTTGGGTAAGGGACCTCCCTGAAGCTCTGTGCTCT	2939
QY	2941	CCAGTATGGTCCCTTTCCCTTCCAGATGCATATGTTTCTTCAGATACAAATAGTG	3000
DB	2940	CCAGTATGGTCCCTTTCCCTTCCAGATGCATATGTTTCTTCAGATACAAATAGTG	2999
QY	3001	ATTCTTAAAAATAACCCAAAAGACAGGCATCCACAGTGTGTGAGCATGAATCAGAGCTGC	3060
DB	3000	ATTCTTAAAAATAACCCAAAAGACAGGCATCCACAGTGTGTGAGCATGAATCAGAGCTGC	3059
QY	3061	ATTGTGTGAGTGGAAATAGTGGGATAAAGTGGATGTACAGAAAGTGGAAATCAACCTC	3120
DB	3060	ATTGTGTGAGTGGAAATAGTGGGATAAAGTGGATGTACAGAAAGTGGAAATCAACCTC	3119
QY	3121	TGCAAAAGCAATCTTTCCTTCTCTGAAGTGTAATAAGAAATACCTCAAGTCTGTGTGTG	3180
DB	3120	TGCAAAAGCAATCTTTCCTTCTCTGAAGTGTAATAAGAAATACCTCAAGTCTGTGTGTG	3179
QY	3181	TGTTGTTACCCAGACTGTCAATCAATAAGACCCAGACTGTCAATGAAAAAATAAAAAA	3240

```

D6 3180 TGGTGGTACCGACACTGTCAATCACTAAAGACCCGACACTCTCAATGAAAAAATAAAAAA 3239
QY 3241 AAAAAAAAAAAAAA 3255
   |||||
D6 3240 AAAAAAAAAAAAAA 3254

RESULT 6
AAFF5254
ID AAFF5254 standard; cDNA; 3253 bp.
XX AC
XX AAFF5254:
XX
XX 29-MAY-2001 (first entry)
XX
DE Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
XX Parkin2: Parkinson's disease; 6q25.2-27; neurodegenerative disease;
XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
XX brain tumour; head trauma; stroke; vascular irregularity;
XX metabolic irregularity; ss.
XX
XX Mus sp.
XX
XX OS
XX Key Location/Qualifiers
XX CDS 129..242
FT FT /tag: a
FT FT /product: "truncated parkin2"
XX
XX EP1081225-A1.
XX
XX C7-MAR-2001.
XX
XX 30-AUG-1999; 39EP-3116766.
XX
XX 30-AUG-1999; 39EP-3116766.
XX
XX (EOP-) BIOFRONTERA PHARM GMBH.
XX
XX Luebbert H;
XX
XX WPI: 2001-212797/22.
XX
XX P-PSDB; AAB67527.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
XX producing a transgenic non-human animal as an animal model for
XX neurodegenerative diseases.
XX
XX Claim 3: Page 10-31; 62pp; English.
XX
XX The present sequence encodes a murine parkin2 polypeptide. The
XX polynucleotide sequence contains mutations, causing Glu38 to be
XX changed to a stop codon, leading to a truncated protein. Mutations or
XX deletions in the parkin2 gene cause Parkinson's disease in humans. The
XX human parkin2 gene is located in gene region 6q25.2-27. Parkin2
XX polypeptides and polynucleotides are useful for analysing
XX neurodegenerative diseases. They are also useful for testing the
XX efficacy of the treatment of a neurodegenerative disease such as
XX Parkinson's disease, Alzheimer's disease, Huntington's disease,
XX amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
XX Pick's disease, Prion disease, and secondary causes including
XX Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
XX stroke, vascular irregularities or metabolic irregularities, associated
XX with a less active or non-active parkin protein.
XX
XX Sequence 3253 bp; 851 A; 816 C; 808 G; 776 T; 0 other;
XX
Query Match: 99.8%; Score 3241; DP 22; Length 3253;
Best Local Similarity 99.9%; Proc. NO. 3;
Matches 3253; Conservative 0; Mismatches 0; Indels 2; Gaps 17;
QV 1 CTCAGCGAGGGGAGGGGAGGAGGGGCTGATGACTAGTAAAGCTGCGTGGTGGG 60

```

1 CTCAGCGAGGGAAGGGGAGGAGGCGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG 60
61 AGGCTCGGGGGGGCGCCAGTSCCGCGTAGTCTCTTCGACCCGCGAGCCACACCGCGC 120
61 ASGCTCGGGGGGGCGCCAGTSCCGCGTAGTCTCTTCGACCCGCGAGCCACACCGCGC 120
121 CGGTGACCATGATAGTGTTCAGGTTCAACTCCAGCTATGCTTCCCGAGTGGAGTCCG 180
121 CGGTGACCATGATAGTGTTCAGGTTCAACTCCAGCTATGCTTCCCGAGTGGAGTCCG 180
181 ATTCTGACACAGAGTCTTCGAGCTCAAGGAAGTGGTGTCTAAGCAGACAGGAGTCCAG 240
181 ATTCTGACACAGAGTCTTCGAGCTCAAGGAAGTGGTGTCTAAGCAGACAGGAGTCCAG 240
238
241 CTGACCACTCGGTGTGATTTTCGCGGGAAGGAGCTTCGGAATCACTGACGTTCCAAA 300
239 CTGACCACTCGGTGTGATTTTCGCGGGAAGGAGCTTCGGAATCACTGACGTTCCAAA 300
301 ACTGTGACCTGGAAACACAGAGTATTTGTACACATAGTACAGAGACCCAGGAGGAATC 360
239 ACTGTGACCTGGAAACACAGAGTATTTGTACACATAGTACAGAGACCCAGGAGGAATC 360
361 ATGAACCAATTCGCTCGAGGGGACGAAACCCGACAGCACTCAGAGGGTCCATATGGG 420
359 ATGAACCAATTCGCTCGAGGGGACGAAACCCGACAGCACTCAGAGGGTCCATATGGG 420
421 AGTCCAGAGAGTTGACACGAGTGGACCTGACAGCCATACCTCCCGGTGGACTCTGG 480
419 AGTCCAGAGAGTTGACACGAGTGGACCTGACAGCCATACCTCCCGGTGGACTCTGG 480
481 GGCTGGGGGTATTCTGGACACAGACAGTAAAGGGATTAGAGAGCCAGGAGTCCAG 540
479 GGCTGGGGGTATTCTGGACACAGACAGTAAAGGGATTAGAGAGCCAGGAGTCCAG 540
541 TTAACCCACCTACAAACAGCTTTTCATCTACTTCAAGGCGCCCTGCCACAAAGTCCAG 600
539 TTAACCCACCTACAAACAGCTTTTCATCTACTTCAAGGCGCCCTGCCACAAAGTCCAG 600
601 CTGGAAGCTCCGAGTTTCAGTGTGGGACCTCCAAACAGCAACCCCTCACCTTGGCCGAG 660
599 CTGGAAGCTCCGAGTTTCAGTGTGGGACCTCCAAACAGCAACCCCTCACCTTGGCCGAG 660
661 GCCATCTTGTGGGACGATCTTTAATTCGAACCGGATAGTGGTGGTGGTGGTGGTGGT 720
659 GCCATCTTGTGGGACGATCTTTAATTCGAACCGGATAGTGGTGGTGGTGGTGGTGGT 720
721 CAGACTGCCCTGGAAACCAAGAGCTGAATTTTCTTTAATGTGGAGCACACCCACCTCAG 780
719 CAGACTGCCCTGGAAACCAAGAGCTGAATTTTCTTTAATGTGGAGCACACCCACCTCAG 780
781 ACAAGGACACGTCGGTAGCTTTGAACCTGTATCAGACAGAGCGGAGGATCCCTTGA 840
779 ACAAGGACACGTCGGTAGCTTTGAACCTGTATCAGACAGAGCGGAGGATCCCTTGA 840
841 TAGCGTGCACAGATGTGAGGAGCCCTGCTGGTCTTCCAGTGTAAACACCGCTCAGTGA 900
839 TAGCGTGCACAGATGTGAGGAGCCCTGCTGGTCTTCCAGTGTAAACACCGCTCAGTGA 900
901 TCTGTTTGGAGTGTTCACCTTGTATGTGTGACAAAGACTCAAGAGTCCGCAATTTTGTCC 960
899 TCTGTTTGGAGTGTTCACCTTGTATGTGTGACAAAGACTCAAGAGTCCGCAATTTTGTCC 960
961 ACGATGCTCAACTTGGCTACTCCCTGCGGTGTGTAGTGGTGTTCGCAATCCCTGATTA 1020
959 ACGATGCTCAACTTGGCTACTCCCTGCGGTGTGTAGTGGTGTTCGCAATCCCTGATTA 1020
1021 AAGAGTCCATCTACCTCAGGATCTCTGGAGAGAGAGTACACTAGGATCCAGCAGTATG 1080
1019 AAGAGTCCATCTACCTCAGGATCTCTGGAGAGAGAGTACACTAGGATCCAGCAGTATG 1080
1081 GGGCGGAGGAATGGGTGCTGCAAAATGGGAGTGTGCTGTGCGCCCGCTCTGGTGGGAG 1140

1079 GSSCCGAGAGGAATSCGTGCTGCAAAATGGGAGGTGTGCTGTGCCCCCGCTCTGCTGTGAG 1138
1141 CTGGAGCTGTACTCTCAACAGGSCCAGAGAAAGTCACTCACTGCGAGGGGGCGAACGCCCTGG 1200
1139 CTGGAGCTGTACTCTCAACAGGSCCAGAGAAAGTCACTGCGAGGGGGCGAACGCCCTGG 1138
1201 GCTCGGGGTTGTTTTCTGCGGGGACTGTAAGGAAGCATACCATGAAGGGGATTCGACT 1260
1199 GCTCGGGGTTGTTTTCTGCGGGGACTGTAAGGAAGCATACCATGAAGGGGATTCGACT 1258
1261 CACTGCTGGAACCCCTCAGAGAGCCACTTCTCAGGCTCAGAGGTGAGCAAAAGAGCGGTG 1320
1259 CACTGCTGGAACCCCTCAGAGAGCCACTTCTCAGGCTCAGAGGTGAGCAAAAGAGCGGTG 1318
1321 ACAGAGCTGCTGAGGAGGAGGCTCCAGGAAGAACATCAAGAAAGACCAACCAAGCTTCTC 1380
1319 ACAGAGCTGCTGAGGAGGAGGCTCCAGGAAGAACATCAAGAAAGACCAACCAAGCTTCTC 1378
1381 CTCGTGCAACGTCGCAATTTGAAAAAACGCGAGGATGTATGCACATGAAGTGTCTCAGC 1440
1379 CTCGTGCAACGTCGCAATTTGAAAAAACGCGAGGATGTATGCACATGAAGTGTCTCAGC 1438
1441 CCGAGCTGCAAGCTGAGTGGTGTGCGGAACTGTGGTGTGTGAGTGGAAAGCGCTCCATGG 1500
1439 CCGAGCTGCAAGCTGAGTGGTGTGCGGAACTGTGGTGTGTGAGTGGAAAGCGCTCCATGG 1498
1501 GAGATCACTGGTGTGACGTGTAGAGAGAGATGTACTTGGCCCTGGAGCAGCAACCTCAA 1560
1499 GAGATCACTGGTGTGACGTGTAGAGAGAGATGTACTTGGCCCTGGAGCAGCAACCTCAA 1558
1561 GGGAACTCCGAAAGTTTCTACTCTTCTTACCATTTCTTTCTGATGATATATAGCA 1620
1559 GGGAACTCCGAAAGTTTCTACTCTTCTTACCATTTCTTTCTGATGATATATAGCA 1618
1621 CATAAATGCGCACACACAAACAGAGCTGAGATACAGAAAGCGCCCTAGATCCTTC 1680
1619 CATAAATGCGCACACACAAACAGAGCTGAGATACAGAAAGCGCCCTAGATCCTTC 1678
1661 TAGGSCACCCACAGAAACCCAGCAGCCCGGTGGSCCGAGGGGAGGAGGACTTTTACG 1740
1679 TAGGSCACCCACAGAAACCCAGCAGCCCGGTGGSCCGAGGGGAGGAGGACTTTTACG 1738
1741 CTCCTGGCTCATCTGATGTGAGAGCTTAGATGAGGCTSCACCTTTGGTTGATCTCTGTA 1800
1739 CTCCTGGCTCATCTGATGTGAGAGCTTAGATGAGGCTSCACCTTTGGTTGATCTCTGTA 1798
1801 GAAGCCATGAGTGGTGGAGTGTTCCTCAGGCTGTTCAGAGCTGTTCAGAGCTGGGTAA 1860
1799 GAAGCCATGAGTGGTGGAGTGTTCCTCAGGCTGTTCAGAGCTGGGTAA 1858
1861 CACTCTGAGGATTTCTCAGAGACACCTTGAGATCTGAGGAACGCTGCTCTCATGTAGTA 1920
1859 CACTCTGAGGATTTCTCAGAGACACCTTGAGATCTGAGGAACGCTGCTCTCATGTAGTA 1918
1921 ATCATCTTCCCAAAGGCGCCCTGCGAGTGTGCAAACTATTTGTTTATCCGCCCAAT 1980
1919 ATCATCTTCCCAAAGGCGCCCTGCGAGTGTGCAAACTATTTGTTTATCCGCCCAAT 1978
1981 CCTATCTTACAAATSGTGTGATGAGATTACAAACCCCTCTGTGTACTTAATCAGCTTATC 2040
1979 CCTATCTTACAAATSGTGTGATGAGATTACAAACCCCTCTGTGTACTTAATCAGCTTATC 2038
2041 ACCCAAGTGAAGACTAGAAAGCTTAATTGATGAGAGCTGCTTAAATCCAGGAGGA 2100
2039 AACCAAGTGAAGACTAGAAAGCTTAATTGATGAGAGCTGCTTAAATCCAGGAGGA 2098
2101 CTCAGAGCGAACTACTCTCGTTCGGTTTCATTTATCTGCAACTTTAGAAAGAAATGATC 2160
2099 CTCAGAGCGAACTACTCTCGTTCGGTTTCATTTATCTGCAACTTTAGAAAGAAATGATC 2158
2161 TTTTTCCTCCCTGAAAAGATAACAAAGTGTGCAATTTGGTTTGGAGTATTTACTGCA 2220
2159 TTTTTCCTCCCTGAAAAGATAACAAAGTGTGCAATTTGGTTTGGAGTATTTACTGCA 2218

QY 121 CCGTGACCATGATAGTGTGTTGTCAGGTTCAATCCAGCTATGCTTCCAGTGGAGGTGG 180
DB |||||
QY 121 CCGTGACCATGATAGTGTGTTGTCAGGTTCAATCCAGCTATGCTTCCAGTGGAGGTGG 180
DB |||||
QY 181 ATTCTGACACAGCATCTTCCAGCTCAAGGAAGTGGTTGCTAAGCGACAGAGGGGTTCCAG 240
DB |||||
QY 181 ATTCTGACACAGCATCTTCCAGCTCAAGGAAGTGGTTGCTAAGCGACAGAGGGGTTCCAG 240
DB |||||
QY 241 CTGACACAGCTCGCTGTGATATTTTCCGGGAGAGAGCTTCGGAATCACTCGAGCTTCAAA 300
DB |||||
QY 241 CTGACACAGCTCGCTGTGATATTTTCCGGGAGAGAGCTTCGGAATCACTCGAGCTTCAAA 300
DB |||||
QY 301 ACTGTGACCTGGGAACAACAGAGTATTGTACACATAGTACAGAGACACAGGAGGAGATC 360
DB |||||
QY 301 ACTGTGACCTGGGAACAACAGAGTATTGTACACATAGTACAGAGACACAGGAGGATC 360
DB |||||
QY 361 ATGAACAAATGATCTGGAGGGAAGAACCCAGAGACACTCAGAGGCTCCATATGGG 420
DB |||||
QY 361 ATGAACAAATGATCTGGAGGGAAGAACCCAGAGACACTCAGAGGCTCCATATGGG 420
DB |||||
QY 421 AGTCCAGAGCTTGACAGAGTGAAGTGAAGAGCTTACCTGAGAGCTTCCGCTGTGG 480
DB |||||
QY 419 AGTCCAGAGCTTGACAGAGTGAAGTGAAGAGCTTACCTGAGAGCTTCCGCTGTGG 478
DB |||||
QY 481 GGCTGGCGGTCACTTCTGGACACAGACAGTAAAGAGGATTCAGAGCAGCAGAGGTCCAG 540
DB |||||
QY 479 GGCTGGCGGTCACTTCTGGACACAGACAGTAAAGAGGATTCAGAGCAGCAGAGGTCCAG 538
DB |||||
QY 541 TTAACCCCACTCAACAGCTTTTTCATCTACTGCAAGGCCCTGCCACAGAGTCCAGC 600
DB |||||
QY 539 TTAACCCCACTCAACAGCTTTTTCATCTACTGCAAGGCCCTGCCACAGAGTCCAGC 598
DB |||||
QY 601 CTGGAAGCTCCAGATTCAGTGTGGACCTGCAACCAAGCAACCCCTCACTTGGGCCAGG 660
DB |||||
QY 599 CTGGAAGCTCCAGATTCAGTGTGGACCTGCAACCAAGCAACCCCTCACTTGGGCCAGG 658
DB |||||
QY 661 GCCCATCTTCTGGAGAGTGTCTTAATTCGAACCGGATGAGTGTGAGTGCAGTCTC 720
DB |||||
QY 659 GCCCATCTTCTGGAGAGTGTCTTAATTCGAACCGGATGAGTGTGAGTGCAGTCTC 718
DB |||||
QY 721 CAGACTGCCCTGGAACACAGAGCTGAATTTTCTTTAAATGTGGAGCACCAACCTCAG 780
DB |||||
QY 719 CAGACTGCCCTGGAACACAGAGCTGAATTTTCTTTAAATGTGGAGCACCAACCTCAG 778
DB |||||
QY 781 ACAAGGACAGCTCGGTAGCTTTGAACCTGATCACAGCAACAGCGCAGCATCCCTTGA 840
DB |||||
QY 779 ACAAGGACAGCTCGGTAGCTTTGAACCTGATCACAGCAACAGCGCAGCATCCCTTGA 838
DB |||||
QY 841 TAGCGTGACAGATGTCAGAGCCCTGCTCTGCTTCCAGTGTAAACACCGCTCAGGTGA 900
DB |||||
QY 839 TAGCGTGACAGATGTCAGAGCCCTGCTCTGCTTCCAGTGTAAACACCGCTCAGGTGA 898
DB |||||
QY 901 TCTGTTTGGACTGTTTCCACTGTTATGTGTCACAAGACTCAACAGTTCGSCAGTTTGTCC 960
DB |||||
QY 899 TCTGTTTGGACTGTTTCCACTGTTATGTGTCACAAGACTCAACAGTTCGSCAGTTTGTCC 958
DB |||||
QY 961 ACGNTGCTCAACTGCTCACTCCCTGCCGTGTAGTGTGCTGCTGCTCAACTCCCTGATTA 1020
DB |||||
QY 959 ACGNTGCTCAACTGCTCACTCCCTGCCGTGTAGTGTGCTGCTGCTCAACTCCCTGATTA 1018
DB |||||
QY 1021 AAGAGCTCCATCACTTCAGGATCTTGGGAAGAGCAGTACACTAGGTACACAGCAGTATG 1080
DB |||||
QY 1019 AAGAGCTCCATCACTTCAGGATCTTGGGAAGAGCAGTACACTAGGTACACAGCAGTATG 1078
DB |||||
QY 1081 GGGCCGAGGAATGCTGCTGGAATGGAGGTGTGTGTGCCCTGCTGCTGCTGCTGAG 1140
DB |||||
QY 1079 GGGCCGAGGAATGCTGCTGGAATGGAGGTGTGTGTGCCCTGCTGCTGCTGCTGAG 1138
DB |||||
QY 1141 CTGAGCTTACCTGAACAGGGCCAGAGGAAGTCACTGCGAGAGGGGGCAAGGGCTGG 1200
DB |||||
QY 1139 CTGAGCTTACCTGAACAGGGCCAGAGGAAGTCACTGCGAGAGGGGGCAAGGGCTGG 1198
DB |||||

QY 1201 GCTGGGGTTTGTCTTCTGCCGGAGCTGAAGGAACATACCATGAAGGGGATTTGGACT 1260
DB |||||
QY 1199 GCTGGGGTTTGTCTTCTGCCGGAGCTGAAGGAACATACCATGAAGGGGATTTGGACT 1258
DB |||||
QY 1261 CACTGTCTGAACCTTCAGGAGCAGCTTCTCAAGCCTTACAGGCTGGAACAAAGAGCGGCTG 1320
DB |||||
QY 1259 CACTGTCTGAACCTTCAGGAGCAGCTTCTCAAGCCTTACAGGCTGGAACAAAGAGCGGCTG 1318
DB |||||
QY 1321 ASCAAGCTGCTCGGAGGAGGCTCCAGGAACCACTCAAGGAACCACTCAAGGAACCACTCAAG 1380
DB |||||
QY 1319 ASCAAGCTGCTCGGAGGAGGCTCCAGGAACCACTCAAGGAACCACTCAAGGAACCACTCAAG 1378
DB |||||
QY 1381 CTGCTGTCGAACCTTGGCAATTTGAAAAAAGGAGGATGTATGCACATGAAGTGTCTCAGC 1440
DB |||||
QY 1379 CTGCTGTCGAACCTTGGCAATTTGAAAAAAGGAGGATGTATGCACATGAAGTGTCTCAGC 1438
DB |||||
QY 1441 CCAAGTGCAGAGCTGGAGTGTGCTGAACTGTGGTGTGTAGTGGGAACCAAGCTTCATGG 1500
DB |||||
QY 1439 CCAAGTGCAGAGCTGGAGTGTGCTGAACTGTGGTGTGTAGTGGGAACCAAGCTTCATGG 1498
DB |||||
QY 1501 GAGATCACTGGTTTGACGTGTAGAGAGAGATGTCACTTGGCCTGAGCAGCAAACTCAA 1560
DB |||||
QY 1499 GAGATCACTGGTTTGACGTGTAGAGAGAGATGTCACTTGGCCTGAGCAGCAAACTCAA 1558
DB |||||
QY 1561 GGGAAACTCCGAAGATTCCTTACCTTCTTAGGCAATTCCTTCTTCTCGATSCATATAAGCA 1620
DB |||||
QY 1559 GGGAAACTCCGAAGATTCCTTACCTTCTTAGGCAATTCCTTCTTCTCGATSCATATAAGCA 1618
DB |||||
QY 1621 CATAAATCGGCAACACACAGCTGCAGATTCAGAAACAGCCCTAGATCTCTTTC 1680
DB |||||
QY 1619 CATAAATCGGCAACACACAGCTGCAGATTCAGAAACAGCCCTAGATCTCTTTC 1678
DB |||||
QY 1681 TAGGCAACCCACAGAAACACACAGCAGCCCTGCTGCTCCAGGGGAGGAGGAGGAGGAGG 1740
DB |||||
QY 1679 TAGGCAACCCACAGAAACACACAGCAGCCCTGCTGCTCCAGGGGAGGAGGAGGAGGAGG 1738
DB |||||
QY 1741 CTCTGGCTCACTCGAATGTGAGAGCTTAGATGAGGCTGCAGCTTGGTTTGGATTTCTGTA 1800
DB |||||
QY 1739 CTCTGGCTCACTCGAATGTGAGAGCTTAGATGAGGCTGCAGCTTGGTTTGGATTTCTGTA 1798
DB |||||
QY 1801 GAAGCCATAGTGAAGTGGGAAGTGTTCAGAGGTGTGTGGCAAGCCCTGGGTGAAGTAA 1860
DB |||||
QY 1799 GAAGCCATAGTGAAGTGGGAAGTGTTCAGAGGTGTGTGGCAAGCCCTGGGTGAAGTAA 1858
DB |||||
QY 1861 CATCTGTGAGATTCAGMAAGCAGACTGAGATCTGAGGAACGCTCTCTCATGTAGTA 1920
DB |||||
QY 1859 CATCTGTGAGATTCAGMAAGCAGACTGAGATCTGAGGAACGCTCTCTCATGTAGTA 1918
DB |||||
QY 1921 ATCACTATTCCCAAGGGGCCCTCCAGTAGTCAAAACTATTGTGCTATCCCCCAAT 1980
DB |||||
QY 1919 ATCACTATTCCCAAGGGGCCCTCCAGTAGTCAAAACTATTGTGCTATCCCCCAAT 1978
DB |||||
QY 1991 CCTATCTTACAAATGGTGTGATGATTAACACCCCTCTGTGTACTAATCAGCTTATC 2040
DB |||||
QY 1979 CCTATCTTACAAATGGTGTGATGATTAACACCCCTCTGTGTACTAATCAGCTTATC 2038
DB |||||
QY 2041 AACCAAGTGAAGACCTTGGGAAGCTTAATGGATGGAGACTGCTTAAATCGGAGGAGGA 2100
DB |||||
QY 2039 AACCAAGTGAAGACCTTGGGAAGCTTAATGGATGGAGACTGCTTAAATCGGAGGAGGA 2098
DB |||||
QY 2101 CTGAGAGCCAAACCTACTTCCCTTGGTTTCATCTATCTCAACTTTTGAAGAAATGATC 2160
DB |||||
QY 2099 CTGAGAGCCAAACCTACTTCCCTTGGTTTCATCTATCTCAACTTTTGAAGAAATGATC 2158
DB |||||
QY 2161 TTTTCTTCCCTCGAAAGATTAACAAAGTCTGGAATTTGGTTGGAGTATTCCTACTGCA 2220
DB |||||
QY 2159 TTTTCTTCCCTCGAAAGATTAACAAAGTCTGGAATTTGGTTGGAGTATTCCTACTGCA 2218
DB |||||
QY 2221 GCTGGAAGTTAGCTTCACTGTGAATTTAAGAGAAAGTGCCTTATAAGGGGGCTTT 2280
DB |||||
QY 2219 GCTGGAAGTTAGCTTCACTGTGAATTTAAGAGAAAGTGCCTTATAAGGGGGCTTT 2278
DB |||||
QY 2281 TTAAGAGACAATCCATGATGTGGTGGGCCAATGCTAAACAAGGGGTCAATAAACAATGT 2340
DB |||||

b	2279	TTATAGAGCAATCCCATGATGCTGGCCAAATGCTATCAACAGGGTCAAGAACAACAATGT	2338
iy	2341	TTATAGAGGAGCATCCCTCGACCATCTGAATGAGAGTATGCTGACCCCTCTCCACCACA	2400
ib	2339	TTATAGAGGAGCATCCCTCGACCATCTGAAATGAGAGTATGCTGACCCCTCTCCACCACA	2398
iy	2401	AGTGGGACACCTCTGCATATCTGCTCCCTCTCTCTGCTTTAAGCCCCAGGAGCCCAT	2460
ib	2399	AGTGGGACACCTCTGCATATCTGCTCCCTCTCTCTGCTTTAAGCCCCAGGAGCCCAT	2458
iy	2461	CCACCAGTGTCTCTCAGACAGGGGAATACACACACCAAGATAGCTTTTCAGATCAAC	2520
ib	2459	CCACCAGTGTCTCTCAGACAGGGGAATACACACACCAAGATAGCTTTTCAGATCAAC	2518
iy	2521	ATGCATCACACTCAAGTGTAAATCTTTCAAGGTTTTCTTTTCTCTTTTCTTTTATT	2580
ib	2519	ATGCATCACACTCAAGTGTAAATCTTTCAAGGTTTTCTTTTCTCTTTTCTTTTATT	2578
iy	2581	TGTTTGTGTTTTGCTTTTTTTTTTTTTTTTTTTTTTGGTGGGCTTACCAAACTTGAG	2640
ib	2579	TGTTTGTGTTTTGCTTTTTTTTTTTTTTTTTTTTTTGGTGGGCTTACCAAACTTGAG	2638
iy	2641	GCCTAGAGCTAAAAATCATATAGAAATGATTTATCTTTGTGTGTGAGGAAAGCCAGCT	2700
ib	2639	GCCTAGAGCTAAAAATCATATAGAAATGATTTATCTTTGTGTGTGAGGAAAGCCAGCT	2698
iy	2701	GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCAGCAGCTCCCAAAAT	2760
ib	2699	GGCCTAAGTTCACACTTTTGTCCAGTGGCCCTAGACTCCACCAGCAGCTCCCAAAAT	2758
iy	2761	GAAAGACCACTCTGCTCAAGCAGCAGTCAGAGTCTGATGTCAACCATCACTATTTTTTT	2820
ib	2759	GAAAGACCACTCTGCTCAAGCAGCAGTCAGAGTCTGATGTCAACCATCACTATTTTTTT	2818
iy	2821	CCATCATGTGCTTGCCTCTGCTCTTCCACACCCGCTGAGCTTAATCGCATTTGGGAG	2880
ib	2819	CCATCATGTGCTTGCCTCTGCTCTTCCACACCCGCTGAGCTTAATCGCATTTGGGAG	2878
iy	2881	CCAGACCAATGTTTGCCTTTCTGCTTTGGGTAAAGGAGCTCCCTCAAGCTCTGTTGGCTCT	2940
ib	2879	CCAGACCAATGTTTGCCTTTCTGCTTTGGGTAAAGGAGCTCCCTCAAGCTCTGTTGGCTCT	2938
iy	2941	CCAGTATGCTCCCTTTTCCCTTCCCTAACAGATGCATATGTTTTCTTCAGAAATACAACTG	3000
ib	2939	CCAGTATGCTCCCTTTTCCCTTCCCTAACAGATGCATATGTTTTCTTCAGAAATACAACTG	2998
iy	3001	ATTCCTTAATAATGCCAAAGACAGCATCCACAGTGTGTGAGCATGAATCACAGCTGC	3060
ib	2999	ATTCCTTAATAATGCCAAAGACAGCATCCACAGTGTGTGAGCATGAATCACAGCTGC	3058
iy	3061	ATTGTGTGAGTGTGAAATAGTGGATAAAAGTGGATGTGAGAGAGTGGAAATCAAACTC	3120
ib	3059	ATTGTGTGAGTGTGAAATAGTGGATAAAAGTGGATGTGAGAGAGTGGAAATCAAACTC	3118
iy	3121	TGCNAGCAATCTTCTCTCTCTGTGAGTGTATTAAGAAATACCTGAAGTCTGTGTGTG	3180
ib	3119	TGCNAGCAATCTTCTCTCTCTGTGAGTGTATTAAGAAATACCTGAAGTCTGTGTGTG	3178
iy	3181	TGTTGGTACCCAGACTGTCAATCAATAAAGACCCAGACTGTCAATTAAGAAAAA	3240
ib	3179	TGTTGGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATTAAGAAAAA	3238
iy	3241	AAAAAAAAAAAAAA 3255	
ib	3239	AAAAAAAAAAAAAA 3253	

RESULT 8

AAF55251

ID AAF55251 standard; cDNA; 3170 BP.

XX

AC AAF55251:

```

XX 29-MAY-2001 (first entry;
XX
XX Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX
XX Parkin2: Parkinson's disease; 6q25.2-27; neurodegenerative disease;
XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
XX brain tumor; head trauma; stroke; vascular irregularity;
XX metabolic irregularity; ss.
XX
XX You sp.
XX
XX Key Location/Qualifiers
XX CDS :29..713
XX F: /tag= a
XX FT /product= "truncated parkin2"
XX
XX PN EP:061225-Al.
XX
XX X 07-MAR-2001.
XX
XX X 30-AUG-1999; 99EP-0116766.
XX
XX X 30-AUG-1999; 99EP-0116766.
XX
XX (EIOF-1 BIOFRONTIERA PHARM GMBH.
XX
XX Luebbert H;
XX
XX X 2001-212797722.
XX P-PSDB: AAB67524.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
XX producing a transgenic non-human animal as an animal model for
XX neurodegenerative diseases.
XX
XX X Claim 3; Page 26-27; 62pp; English.
XX
XX The present sequence encodes a murine parkin2 polypeptide. The
XX polynucleotide sequence contains a frameshift mutation, leading to a
XX truncated protein. Mutations or deletions in the parkin2 gene cause
XX Parkinson's disease in humans. The human parkin2 gene is located in
XX gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are
XX useful for analysing neurodegenerative diseases. They are also useful
XX for testing the efficacy of the treatment of a neurodegenerative disease,
XX such as Parkinson's disease, Alzheimer's disease, Huntington's disease,
XX amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
XX Pick's disease, Prion disease, and secondary causes inducing
XX Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
XX stroke, vascular irregularities or metabolic irregularities, associated
XX with a less active or non-active parkin protein.
XX
XX X Sequence 3170 BP; 834 A; 735 C; 784 G; 757 T; 3 other;
XX
XX Query Match: 94.5%; Score 3075; DB 22; Length 3170;
XX Best Local Similarity 97.4%; Pred. No. C;
XX Matches 3170; Conservative 0; Mismatches 0; Indels 85; Gaps 1
XX
XX 1 CTCACGAGGGGAGGGGGAGGCGGTGATGACTAACTGACACAGAAACGGTGGG 60
XX :
XX 2 CTCACGAGGGGAGGGGGAGGAGAGGCTGATGACTAACTGACAGAAACGGTGGG 60
XX
XX 61 AGGCTCGGGCGGGGCCAGTGCCTCGGTAGTGCCTTCGACCGGAGCACCACCGGCC 120
XX
XX 61 AGGCTCGGGCGGGGCCAGTGCCTCGGTAGTGCCTTCGACCGGAGCACCACCGGCC 120
XX
XX 121 CGGTGACCATGATAGTGTTCCTCAGGTTCCACTCCAGCTAGCTTCCCGAGTGCAGTGC 190
XX
XX 121 CGGTGACCATGATAGTGTTCCTCAGGTTCCACTCCAGCTAGCTTCCCGAGTGCAGTGC 190
XX
XX 161 ATTCTGACCAAGCATGTTGACGATCAAGGAAGTGGTTTGTAGGACACAGGGGTTCAG 240
XX

```

```

RESULT 8
AAF55251
ID AAF55251 standard; cDNA; 3,170 bp.
XX
AC AAF55251;

```

DB 181 ATTCTGACACCAAGCATCTTTGCGAGCTCAAGAAAGTGGTTGCTTAAGCGACAGGGGTTCCAG 240
 QY 241 CTGACGAGCTGGCTGTGATTTTTCGGGGAAGGAGCTTCCGAATCACCTGACGCTTCAAA 303
 DB 241 CTGACGAGCTGGCTGTGATTTTTCGGGGAAGGAGCTTCCGAATCACCTGACGCTTCAAA 303
 QY 301 ACTGTGACTGTGGAACCAACAGAGATTTGTACACATAGTACAGAGACCAAGGAGAGAGTC 360
 DB 301 ACTGTGACTGTGGAACCAACAGAGATTTGTACACATAGTACAGAGACCAAGGAGAGAGTC 360
 QY 361 ATGAACAAATGCACTGTGAGGGGAGCAACCCACAGACACCTTCAGAGGGTCCATATGGG 423
 DB 361 ATGAACAAATGCACTGTGAGGGGAGCAACCCACAGACACCTTCAGAGGGTCCATATGGG 420
 QY 421 AGTCCAGGAGCTTGACACAGAGTGGACTGAGCAGCATACCTGCGGTGACTCTGTGG 480
 DB 421 AGTCCAGGAGCTTGACACAGAGTGGACTGAGCAGCATACCTGCGGTGACTCTGTGG 480
 QY 481 GGTGGCGGTCACTTCGGACAGACAGTAAAGAGGATTCAGACAGCCACAGAGTCCAG 540
 DB 481 GGTGGCGGTCACTTCGGACAGACAGTAAAGAGGATTCAGACAGCCACAGAGTCCAG 540
 QY 541 TTAACCCACCTACAAACAGCTTTTTCATCTACTGTCAAAAGGCCCTGCCAAGTCCAGC 600
 DB 541 TTAACCCACCTACAAACAGCTTTTTCATCTACTGTCAAAAGGCCCTGCCAAGTCCAGC 600
 QY 601 CTGAAAGGCTCCGAGTTCAGTGTGGCAGCTGCAAAACAGCAGCTCACCTGSCCCAGG 660
 DB 601 CTGAAAGGCTCCGAGTTCAGTGTGGCAGCTGCAAAACAGCAGCTCACCTGSCCCAGG 660
 QY 661 GCGCATCTTCTGGGAGCATGTCTTAATTTCCAAACCGGATGAGTGGTGAATGCCAGTCTC 720
 DB 661 ----- 720
 QY 721 CAGACTGCCCTGGACACGAGCTGAAATTTCTTAAATGTGGACACACCAACCTCAG 780
 DB 659 ----- 659
 QY 781 ACAAGGACAGCTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTTGA 840
 DB 696 ACAAGGACAGCTCGGTAGCTTTGAACCTGATCACCAGCAACAGGCGCAGCATCCCTTGA 755
 QY 841 TAGCGTGACAGATGTGAGAGCCCTGCTGGTCTTCAGTGTAAACACCGTCAGCTGA 900
 DB 756 TAGCGTGACAGATGTGAGAGCCCTGCTGGTCTTCAGTGTAAACACCGTCAGCTGA 815
 QY 901 TCTGTGTGAGCTGTTCCACTGTGTATGTGTACAGACTCAACGATCGGAGTTGTCC 960
 DB 816 TCTGTGTGAGCTGTTCCACTGTGTATGTGTACAGACTCAACGATCGGAGTTGTCC 875
 QY 961 ACAGTGTCAACTGGCTACTCCCTGCGCTGTGAGCTGGCTGTCCAACTCCCTGATTA 1020
 DB 976 ACAGTGTCAACTGGCTACTCCCTGCGCTGTGAGCTGGCTGTCCAACTCCCTGATTA 935
 QY 1021 AAGAGCTCCATCACTTCAGGATCTTTGAGAGAGCAGTACACTAGGTACCAGCAGTATG 1080
 DB 936 AAGAGCTCCATCACTTCAGGATCTTTGAGAGAGCAGTACACTAGGTACCAGCAGTATG 995
 QY 1081 GSGCCAGGAATGCTGTCCAAATGGGAGGTGTGCTGTGCCCGCTGCTGGCTGTGGAG 1140
 DB 996 GSGCCAGGAATGCTGTCCAAATGGGAGGTGTGCTGTGCCCGCTGCTGGCTGTGGAG 1055
 QY 1141 CTGAGCTGCTACTGGAACAGGGCCAGAGAAAGTCACTTGGCAAGGGGCAACGGCCTGG 1200
 DB 1056 CTGAGCTGCTACTGGAACAGGGCCAGAGAAAGTCACTTGGCAAGGGGCAACGGCCTGG 1115
 QY 1201 GCTCGGGTGTGTTTCTGCGGAGCTGTAAAGAGCATACCATGAAGGGGATTCGACT 1260
 DB 1116 GCTCGGGTGTGTTTCTGCGGAGCTGTAAAGAGCATACCATGAAGGGGATTCGACT 1175
 QY 1261 CACTGCTCGAGACCTCAGGAGCCACTTCTCAGGCTTACAGGGTGGACAAAGAGCGCTG 1320
 DB 1176 CACTGCTCGAGACCTCAGGAGCCACTTCTCAGGCTTACAGGGTGGACAAAGAGCGCTG 1235

QY 1321 AGCAAGCTTCGCTGGGAGGAGCCTCCAAAGGAACCACTCAAGAAAGACCAAGCCTTTGTC 1380
 DB 1236 AGCAAGCTTCGCTGGGAGGAGCCTCCAAAGGAACCACTCAAGAAAGACCAAGCCTTTGTC 1295
 QY 1381 CTCGCTGCAATGCTGCCAATTTGAANAANAAGAGAGATGTATGGACATGAAGTGTCTTCAGC 1440
 DB 1236 CTCGCTGCAATGCTGCCAATTTGAANAANAAGAGAGATGTATGGACATGAAGTGTCTTCAGC 1355
 QY 1441 GCGAGTCCAAAGCTGAGTGGTGTGGAACTGTGCTGTGAGTGGAAACCGAGGCTGCATGG 1500
 DB 1356 GCGAGTCCAAAGCTGAGTGGTGTGGAACTGTGCTGTGAGTGGAAACCGAGGCTGCATGG 1415
 QY 1501 GAGATCACTGTTTGAAGCTGTAGAGAGAGATGTCAATTTGGCCCTGTGAGCGCAACATGAA 1560
 DB 1416 GAGATCACTGTTTGAAGCTGTAGAGAGAGATGTCAATTTGGCCCTGTGAGCGCAACATGAA 1475
 QY 1561 GGGAAATCTCGAAGATTCCTAGCTTCTTAGCCATTTCTTCTCGATGCATATAAGCA 1620
 DB 1476 GGGAAATCTCGAAGATTCCTAGCTTCTTAGCCATTTCTTCTCGATGCATATAAGCA 1535
 QY 1621 CATAAATGCGGACACACAAACACAGGCTGCAGATTTAGAGAGAGCGCCCTAGATCCTTTC 1680
 DB 1536 CATAAATGCGGACACACAAACACAGGCTGCAGATTTAGAGAGAGCGCCCTAGATCCTTTC 1595
 QY 1681 TAGGGCACCCACACAAACACAGACCCCTGGCTCCAGGGGAGGAGGACACTTTCAGC 1740
 DB 1596 TAGGGCACCCACACAAACACAGACCCCTGGCTCCAGGGGAGGAGGACACTTTCAGC 1655
 QY 1741 CTTGGCTCACTCGAAATGTGAGAGCTTAGATGAGGCTGACCTTTGGTTTGGATTTGTGA 1800
 DB 1656 CTTGGCTCACTCGAAATGTGAGAGCTTAGATGAGGCTGACCTTTGGTTTGGATTTGTGA 1715
 QY 1801 GAAGGCATGAGTGGAGTGGGAAGTGTTCACAGGCTGTGTCACACCCCTGGGTAAAGTAA 1860
 DB 1716 GAAGGCATGAGTGGAGTGGGAAGTGTTCACAGGCTGTGTCACACCCCTGGGTAAAGTAA 1775
 QY 1861 CACCTCTGAGGATTTCTGAGAGGACACTTGAGATCTGAGGAGACGCTGCTCTGTATGTA 1920
 DB 1776 CACCTCTGAGGATTTCTGAGAGGACACTTGAGATCTGAGGAGACGCTGCTCTGTATGTA 1835
 QY 1921 ATCATCTATTCGCAAGGCGCCCTGGCAGTAGTCAAAACTATTTGTTTATCCGCCCAAA 1980
 DB 1836 ATCATCTATTCGCAAGGCGCCCTGGCAGTAGTCAAAACTATTTGTTTATCCGCCCAAA 1895
 QY 1981 CTTATCTTTAATAATGGTGTGATGAGATTACAAACCCCTGTGTGTACTATACGCTTAIC 2040
 DB 1896 CTTATCTTTAATAATGGTGTGATGAGATTACAAACCCCTGTGTGTACTATACGCTTAIC 1955
 QY 2041 AACCAAGTGAGAACTAGGAAGCTAAATGGATGGGAGACTGCTTAAATCGCAGGAGGA 2100
 DB 1956 AACCAAGTGAGAACTAGGAAGCTAAATGGATGGGAGACTGCTTAAATCGCAGGAGGA 2015
 QY 2101 CTCAGAAACCAAACTACTTCCGTTCTGTTTCAATTTCTGAACTTTTAGAAGAAATGATC 2160
 DB 2016 CTCAGAAACCAAACTACTTCCGTTCTGTTTCAATTTCTGAACTTTAGAAGAAATGATC 2075
 QY 2161 TTTTTCCTCCCTGAAAGATACAAAGCTGGAATTTGGTATTTGGAGTATTCCTACTGCA 2220
 DB 2076 TTTTTCCTCCCTGAAAGATACAAAGCTGGAATTTGGTATTTGGAGTATTCCTACTGCA 2135
 QY 2221 GCTGSAAGTGTAGCTTCACTGTGAAATTTAACAGASAAAGTGCCTATAAAGGGGGCTTT 2280
 DB 2136 GCTGSAAGTGTAGCTTCACTGTGAAATTTAACAGASAAAGTGCCTATAAAGGGGGCTTT 2195
 QY 2281 TTAAGAGCAATTCGACATGCTGCGGCAATGCTTAACAGAGGCTGAAGAAACACATGT 2340
 DB 2196 TTAAGAGCAATTCGACATGCTGCGGCAATGCTTAACAGAGGCTGAAGAAACACATGT 2255
 QY 2341 TTAAGAGGAGCAATTCGCTGACCATGTGAATGAGAGATGCTGCTGACCCCTTCCACACA 2400
 DB 2256 TTAAGAGGAGCAATTCGCTGACCATGTGAATGAGAGATGCTGCTGACCCCTTCCACACA 2315

[illegible]

2717 GCACGACAAATGTTTCTGCTTCTGCTTCTGGTAAAGGAGCTCCCTGAAGCTCTGGCTC 2776
 2940 TCACGATGTCCTCTTCTCTCTCTTAACAGATCATATGTTTCTTTCAGAAATACAAATAGT 2999
 2777 TCACGATGTCCTCTTCTCTCTCTTAACAGATCATATGTTTCTTTCAGAAATACAAATAGT 2836
 3000 GATCTTAAATTAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCACAGCCTG 3059
 2837 GATCTTAAATTAACCCAAAGACAGGATCCACAGTGTGTGAGCATGAATCACAGCCTG 2896
 3260 CATCTGTGAGTGTGATGAGTGGATTAAGTGGATGTGAGAGAGTGGAAATCAAACT 3119
 2897 CATCTGTGAGTGTGATGAGTGGATTAAGTGGATGTGAGAGAGTGGAAATCAAACT 2956
 3120 CTGCAAAAGCAATCTTCTCTCTTCTGTAAGTGTATTAAAGAAATACCTGAAGTCTGTGT 3179
 2957 CTGCAAAAGCAATCTTCTCTCTTCTGTAAGTGTATTAAAGAAATACCTGAAGTCTGTGT 3016
 3180 GTGGTGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATGAATAAATAAATAA 3239
 3017 GTGGTGTACCCAGACTGTCAATCAATTAAGACCCAGACTGTCAATGAATAAATAAATAA 3076
 3240 AAAAAAAAAAAAAA 3255
 3077 AAAAAAAAAAAAAA 3092

RESULT 11

AAF55253
 ID AAF55253 standard; cDNA; 3043 BP.

AC AAF55253;

DT 29-MAY-2001 (first entry)

XX Nucleotide sequence of murine parkin2 with a frameshift mutation.

XX Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 XX brain tumour; head trauma; stroke; vascular irregularity;
 XX metabolic irregularity; ss.

XX Mus sp.

FH Key Location/Qualifiers
 FT CDS 129..1019
 FT /*tag= a
 FT /product= "truncated parkin2"

XX EPI081225-A1.

XX 07-MAR-2001.

XX 30-AUG-1999; 99EP-0116766.

XX 30-AUG-1999; 99EP-0116766.

XX (BIOF-) BIOFRONTIERA PHARM GYBHI.

XX Luebbert H;

XX WPI: 2001-242797/22.

XX P-PSDB; AAB67526.

XX New polynucleotides encoding mouse parkin2 protein, useful for
 XX producing a transgenic non-human animal as an animal model for
 XX neurodegenerative diseases

XX Claim 3; Page 29-30; 62pp; English.

XX The present sequence encodes a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains a frameshift mutation, leading to a

CC truncated protein. Mutations or deletions in the parkin2 gene cause
 CC Parkinson's disease in humans. The human parkin2 gene is located in
 CC gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are
 CC useful for analysing neurodegenerative diseases. They are also useful
 CC for testing the efficacy of the treatment of a neurodegenerative disease
 CC such as Parkinson's disease, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 CC Pick's disease, Prion disease, and secondary causes inducing
 CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
 CC stroke, vascular irregularities or metabolic irregularities, associated
 CC with a less active or non-active parkin protein.

XX Sequence 3043 BP; 806 A; 763 C; 737 G; 737 T; 0 other:

Query Match 96.7%; Score 2821; DB 22; Length 3043;

Best Local Similarity 93.5%; Pred. No. 0;

Matches 3043; Conservative 0; Mismatches 3; Indels 212; Gaps 1;

QY 1 CTGAGCAGGGGAGGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG 60

DB 1 CTGAGCAGGGGAGGGGAGGAGGCTGGATGACTAAACCTGACAGAAACGCTGGTGGG 60

QY 61 AGGCTCGGGGGGGGGCAGTGGCCGGGTAGGTCTCTCTGGACCCGAGCCACCCCGGC 120

DB 61 AGGCTCGGGGGGGGGCAGTGGCCGGGTAGGTCTCTCTGGACCCGAGCCACCCCGGC 120

QY 121 CGGTGACCATGATAGTGTCTTGTGAGGTCAACTCCAGCTATGGCTTCCAGTGGAGTGG 180

DB 121 CGGTGACCATGATAGTGTCTTGTGAGGTCAACTCCAGCTATGGCTTCCAGTGGAGTGG 180

QY 181 ATTCTGACACGACATCTTCACACTCAAGGAGTGTGTGCTAGGACACAGGGGTCCAG 240

DB 181 ATTCTGACACGACATCTTCACACTCAAGGAGTGTGTGCTAGGACACAGGGGTCCAG 240

QY 241 CTGACACAGCTGGCTGTGATTTTCCCGGGAAGAGTTCCGAATCACTGACGGTCAAA 300

DB 241 CTGACACAGCTGGCTGTGATTTTCCCGGGAAGAGTTCCGAATCACTGACGGTCAAA 300

QY 301 ACTGTGACCTGGAAACACAGATTTTACACATAGTACAGAGACACGAGGAGAGTC 360

DB 301 ACTGTGACCTGGAAACACAGATTTTACACATAGTACAGAGACACGAGGAGAGTC 360

QY 361 ATGAAACAAATGCTGTGGGGGAGTGAACCCGAGAGCTTCAGAGGGGTTCATAGGG 420

DB 361 ATGAAACAAATGCTGTGGGGGAGTGAACCCGAGAGCTTCAGAGGGGTTCATAGGG 420

QY 421 AGTCCAGGAGCTTACACAGAGTGGACTGAGAGCGATACCTGCGGGTGGACTCTGTGG 480

DB 421 AGTCCAGGAGCTTACACAGAGTGGACTGAGAGCGATACCTGCGGGTGGACTCTGTGG 480

QY 481 GGTGGGGTTCATCTGGACACAGAGTAAAGAGGATTCAGAAAGCCAGGCTCCAG 540

DB 481 GGTGGGGTTCATCTGGACACAGAGTAAAGAGGATTCAGAAAGCCAGGCTCCAG 540

QY 541 TTAACCCACCTACACAGCTTTTTCATCTGCAAGAGCCCTGCCACAGAGTCCAG 600

DB 541 TTAACCCACCTACACAGCTTTTTCATCTGCAAGAGCCCTGCCACAGAGTCCAG 600

QY 601 CTGAAAGCTCCGAGTTCAGTGTGGACCTGCAAAACAACTCACTTGGCCAGG 660

DB 601 CTGAAAGCTCCGAGTTCAGTGTGGACCTGCAAAACAACTCACTTGGCCAGG 660

QY 661 GCGCATCTTCTGGAGAGTGTATTATTCGAAGAGGATGAGTGGTGGATGCCATCTC 720

DB 661 GCGCATCTTCTGGAGAGTGTATTATTCGAAGAGGATGAGTGGTGGATGCCATCTC 720

QY 721 CAGACTCCCTGGAACACAGAGTGAATTTTCTTTAAATGTGAGACACCCCACTCAG 780

DB 721 CAGACTCCCTGGAACACAGAGTGAATTTTCTTTAAATGTGAGACACCCCACTCAG 780

QY 781 ACAAGGACAGTCCGTAGCTTTGAACCTGATCACCAGCAAGGGCCAGCATCCCTGCA 840

DB 781 ACAAGGACAGTCCGTAGCTTTGAACCTGATCACCAGCAAGGGCCAGCATCCCTGCA 840

QY	841	TAGGTCACAGATGTCAGAGCCCTGTCTGTGCTTCCAGTGTAAACACCGTCACTGA	900	1921	ATCATCTATTCCAAAGGGCCCTGCGAGTAGTCAAAACTATTGTTATCCCCCAAT	1980
DB	841	TAGCGTSCACAGATGTCAGAGCCCTGTCTGTGCTTCCAGTGTAAACACCGTCACTGA	900	1922	ATCATCTATTCCAAAGGGCCCTGCGAGTAGTCAAAACTATTGTTATCCCCCAAT	1981
QY	901	TCTGTGGAGCTGTTTCCACTTCTATTGTTGTCACAGACTCAACGATCGGAGTTTGTCC	960	1981	CCATCTTTACAAATGGTGTGATGAGATTACAGCCCTCTGTGACTAATCAGCTTATC	2040
DB	901	TCTGTGGAGCTGTTTCCACTTCTATTGTTGTCACAGACTCAACGATCGGAGTTTGTCC	960	1982	CCATCTTTACAAATGGTGTGATGAGATTACAGCCCTCTGTGACTAATCAGCTTATC	2041
QY	961	ACGATGCTCAACTTGGCTACTCCCTCCCGTGTAGCTGGCTGTCCCACTCCCTGATTA	1020	2041	AAACAGTGGAGAACCTAGAAAGCTAAATGGATGGAGACTGCTTTAAATCCGAGGAGA	2100
DB	961	ACGATGCTCAACTTGGCTACTCCCTCCCGTGTAGCTGGCTGTCCCACTCCCTGATTA	1020	2042	AAACAGTGGAGAACCTAGAAAGCTAAATGGATGGAGACTGCTTTAAATCCGAGGAGA	2101
QY	1021	AAGAGCTCCATCACTTCAGGATCCTTGGAGAGAGCAGTACACTAGGTACAGAGATAG	1080	2101	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2160
DB	996	-----	995	2102	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2161
QY	1081	GGCCGAGGAATCGTGTCAAAATGGAGGTGTGCTGTGCCCGCTCTGTGTGTGAG	1140	2161	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2162
DB	996	-----	995	2162	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2163
QY	1141	CTGACTGTCTACCTGACAGGGCCAGAGGAAGTCACTGCGAAGGGGCAACGGCTGG	1200	2163	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2164
DB	996	-----	995	2164	CTCAGAGSCAAACCTACTCTCGTTCGTTTCATTATCTGCAACTTTAGAAAGAAATGATC	2165
QY	1201	GCTGCGGGTGTGTTTCTGCGGGACTGTAAAGGAATACCATGAAGGGGATTCGACT	1260	2221	GGCTGGAGGTTTAGCTTTCATCTGTGTGTAATTTACAGAGAAAGTCCCTATAAAGGGGGCTT	2280
DB	996	-----	995	2222	GGCTGGAGGTTTAGCTTTCATCTGTGTGTAATTTACAGAGAAAGTCCCTATAAAGGGGGCTT	2281
QY	1261	CAGTCTCGAAGCTCTAGAGGACACTTCTCAGGCTTACAGGCTGACAAAGAGCGGTG	1320	2281	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2340
DB	996	-----	995	2282	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2341
QY	1321	AGCAAGCTGCTGGAGGAGGCTTCAAGGAAACCATCAAGGAAGACCAACAGCCCTGTC	1380	2341	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2342
DB	1109	AGCAAGCTGCTGGAGGAGGCTTCAAGGAAACCATCAAGGAAGACCAACAGCCCTGTC	1168	2342	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2343
QY	1381	CTGCTGCAAGTGCCTAATTTGAAAAAAGCGAGATGTTATGCAATGAAGTGTCTCAGC	1440	2343	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2344
DB	1169	CTGCTGCAAGTGCCTAATTTGAAAAAAGCGAGATGTTATGCAATGAAGTGTCTCAGC	1228	2344	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2345
QY	1441	CCAGTCCAGCTGGAGTGTGTGAACTGTGGCTGTGAGTGGAGCGAGCTCATGG	1500	2345	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2346
DB	1229	CCAGTCCAGCTGGAGTGTGTGAACTGTGGCTGTGAGTGGAGCGAGCTCATGG	1288	2346	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2347
QY	1501	GAGATCACTGTTGACGTGTAGAGAGATGTCACTTGGCCCTGGAGCACAACCTCAA	1560	2347	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2348
DB	1289	GAGATCACTGTTGACGTGTAGAGAGATGTCACTTGGCCCTGGAGCACAACCTCAA	1348	2348	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2349
QY	1561	GGGAACTCCGAGATTCCTACCTTCTTACGATTTCTTCTCGATGCATATAAGCA	1620	2349	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2350
DB	1349	GGGAACTCCGAGATTCCTACCTTCTTACGATTTCTTCTCGATGCATATAAGCA	1408	2350	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2351
QY	1621	CATAATGCGCACACACAAACACAGCTCGAGATTACAGAGACGCCCTAGATCCTTC	1680	2351	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2352
DB	1409	CATAATGCGCACACACAAACACAGCTCGAGATTACAGAGACGCCCTAGATCCTTC	1468	2352	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2353
QY	1681	TAGGGCACACAGAAACACAGACACCGCTGGCCCGAGGGGGAGGAGCACTTTCAGC	1740	2353	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2354
DB	1469	TAGGGCACACAGAAACACAGACACCGCTGGCCCGAGGGGGAGGAGCACTTTCAGC	1528	2354	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2355
QY	1741	CTGTGGCTCACTCGAATGTACAGAGCTTAGATGAGGTGCACCTTTGGTTGGATTTCTGA	1800	2355	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2356
DB	1529	CTGTGGCTCACTCGAATGTACAGAGCTTAGATGAGGTGCACCTTTGGTTGGATTTCTGA	1588	2356	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2357
QY	1801	GAGGCCATGAGTGGAGTGGAGTGTTCAGAGGTGTGGCAGCCCTGGGTAAATGA	1860	2357	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2358
DB	1589	GAGGCCATGAGTGGAGTGGAGTGTTCAGAGGTGTGGCAGCCCTGGGTAAATGA	1648	2358	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2359
QY	1861	CACCTCTGAGGATTTCTCAGAGACACACTTTCAGATCTGAGGACCGCTCTCATCTAGTA	1920	2359	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2360
DB	1649	CACCTCTGAGGATTTCTCAGAGACACACTTTCAGATCTGAGGACCGCTCTCATCTAGTA	1798	2360	TTAAGAGACAAATCCCATGATGCTGCGCAATGCTTAACAGAGAAAGTCCCTATAAAGGGGGCTT	2361

b 660 TCGTTTGGACCTGTTCCACTGTGTATGTGTGCACAAAGACTCAACGATCGGAGTTTGTCC 719
y 961 ACATGCTCAACTTGGCTACTCCCTCCGCTGTGTAGCTGGCTGTGCCAATCCCTGATTA 1020
b 720 ACATGCTCAACTTGGCTACTCCCTCCGCTGTGTAGCTGGCTGTGCCAATCCCTGATTA 779
y 1021 AAGAGTCCATCACTTCAAGATCCCTGGAGAAGAGCAGTACACTAGGTACACAGCAGTATG 1080
b 780 AAGAGTCCATCACTTCAAGATCCCTGGAGAAGAGCAGTACACTAGGTACACAGCAGTATG 839
y 1081 GGCCCGAGGAATCGCTGTGCAATGGAGGTGTGCTGTGCCCTGCTCTGGCTGTGGAG 1140
b 840 GGCCCGAGGAATCGCTGTGCAATGGAGGTGTGCTGTGCCCTGCTCTGGCTGTGGAG 899
y 1141 CTGAGCTGTACCTGAACAGGGCCGAGGAAGTACCTGTCCGAAGGGGCGACGSCCTGG 1200
b 900 CTGAGCTGTACCTGAACAGGGCCGAGGAAGTACCTGTCCGAAGGGGCGACGSCCTGG 959
y 1201 GCTCGGGTTTGTCTTCTGCGGAGCTGTAAAGGAAGCATACCATGAAGGGATTCGCAAT 1260
b 960 GCTCGGGTTTGTCTTCTGCGGAGCTGTAAAGGAAGCATACCATGAAGGGATTCGCAAT 1019
y 1261 CACTGCTCAACCTCAGGACCACTTCTCAGGCTTACAGGCTGAGGCTGACAAAGAGCGGTG 1320
b 1020 CACTGCTCAACCTCAGGACCACTTCTCAGGCTTACAGGCTGAGGCTGACAAAGAGCGGTG 1079
y 1321 AGCAAGCTCGCTGGAGGAGGCTCCAAAGGAACCATCAAGGAAGCACCACCAAGCCTTCTC 1380
b 1080 AGCAAGCTCGCTGGAGGAGGCTCCAAAGGAACCATCAAGGAAGCACCACCAAGCCTTCTC 1139
y 1381 CTGCTGCAACCTGCCAATGAAAGGAAGGAGGATGTATGACATGAAGTGTCTTCAGC 1440
b 1140 CTGCTGCAACCTGCCAATGAAAGGAAGGAGGATGTATGACATGAAGTGTCTTCAGC 1199
y 1441 CCCAGTCAAGCTGGAGTGTGTGCACTGTGGCTGTGAGTGGAAACGAGGCTGCATGG 1500
b 1200 CCCAGTCAAGCTGGAGTGTGTGCACTGTGGCTGTGAGTGGAAACGAGGCTGCATGG 1259
y 1501 GAGATCACTGGTTTGAGCTGTAGAGAGAGATGTCACTTGGCCCTCGAGCGCAACCTCAA 1560
b 1260 GAGATCACTGGTTTGAGCTGTAGAGAGAGATGTCACTTGGCCCTCGAGCGCAACCTCAA 1319
y 1561 GGAAGACTCGAAGATTCCCTACCTCTTAGCCATCTCTTCTCGATGCAATTAAGCA 1620
b 1320 GGAAGACTCGAAGATTCCCTACCTCTTAGCCATCTCTTCTCGATGCAATTAAGCA 1379
y 1621 CATAAATGCGCACACAAACACAGGCTCGAGATTACAGAGAGCGCCCTAGATCCTTC 1680
b 1380 CATAAATGCGCACACAAACACAGGCTCGAGATTACAGAGAGCGCCCTAGATCCTTC 1439
y 1681 TAGGGCACCCAGAAACACAGACCCGCTGGCCCGAGGGGAGGAGGACCTTTCAGC 1740
b 1440 TAGGGCACCCAGAAACACAGACCCGCTGGCCCGAGGGGAGGAGGACCTTTCAGC 1499
y 1741 CTCTGGCTCACTCGAATGTAGATGAGGCTGCACCTTTTGGTTTGGATCTGTGA 1800
b 1500 CTCTGGCTCACTCGAATGTAGATGAGGCTGCACCTTTTGGTTTGGATCTGTGA 1559
y 1801 GAAGCCATGAGTGGAGTGGAGTGTTCAGAGGTTGTGCCAGCCCTGGGTAACTAA 1860
b 1560 GAAGCCATGAGTGGAGTGGAGTGTTCAGAGGTTGTGCCAGCCCTGGGTAACTAA 1619
y 1861 CACTCTGAGATTCTCAGAAACACACITGAGATCTGAGGAAGCGTGTCTCATGTAGTA 1920
b 1620 CACTCTGAGATTCTCAGAAACACACITGAGATCTGAGGAAGCGTGTCTCATGTAGTA 1679
y 1921 ATCATCTATTCCCAAGGSCCCCTCGAGTAGTCAAAACTATTTTTATCCCCCAAAAT 1980
b 1680 ATCATCTATTCCCAAGGSCCCCTCGAGTAGTCAAAACTATTTTTATCCCCCAAAAT 1739
y 1981 CCTATCTTTACAAATGGTGTGATGAGATTACACCCCTCTGTACTAATCAGCTTATC 2040
b 1740 CCTATCTTTACAAATGGTGTGATGAGATTACACCCCTCTGTACTAATCAGCTTATC 1799

Qy 2041 AACCAAGTCAGAAACCTAGAAAGCTAAATGGATGGAGACTGCTTAAATCGCAGGAGGA 2100
Db 1800 AACCAAGTCAGAAACCTAGAAAGCTAAATGGATGGAGACTGCTTAAATCGCAGGAGGA 1859
Qy 2101 CTCAGAAAGCAACCTACTCCGTTCCGTTTCATTCCTGCAACTTCAGAAAGAAATGATC 2160
Db 1860 CTCAGAAAGCAACCTACTCCGTTCCGTTTCATTCCTGCAACTTCAGAAAGAAATGATC 1919
Qy 2161 TTTTTCCTCCCTGAAAGATAACAAAGTGTGCAATTTGGTTCGAGTATTCCTACTGCA 2220
Db 1920 TTTTTCCTCCCTGAAAGATAACAAAGTGTGCAATTTGGTTCGAGTATTCCTACTGCA 1979
Qy 2221 GCTGGAAGTGTAGTTCGACTGTGATTTACAGAGAAAGTGCCCTATAAGGGGGCTTT 2280
Db 1980 GCTGGAAGTGTAGTTCGACTGTGATTTACAGAGAAAGTGCCCTATAAGGGGGCTTT 2039
Qy 2281 TTAAGAGACAATCCCATGATGCTGCGCAATGCTAAACAACAGGGTCAAGAAACACAATGT 2340
Db 2040 TTAAGAGACAATCCCATGATGCTGCGCAATGCTAAACAACAGGGTCAAGAAACACAATGT 2099
Qy 2341 TTATAGAGAGAGAGTCCCTCGACATCTGAATGAGAGTATGCTGACCCCTCCACACA 2400
Db 2100 TTATAGAGAGAGAGTCCCTCGACATCTGAATGAGAGTATGCTGACCCCTCCACACA 2159
Qy 2401 AGTGGGAGACCTCTGCATATCTGCTCCCTCTCTCTGCTTTAAAGCCCGAGGAGCCCAT 2460
Db 2160 AGTGGGAGACCTCTGCATATCTGCTCCCTCTCTCTGCTTTAAAGCCCGAGGAGCCCAT 2219
Qy 2461 CCACCCAGTGTCTTACAGAGAGGGCAATACACACACACAGATAGGCTTCAGATCAAC 2520
Db 2220 CCACCCAGTGTCTTACAGAGAGGGCAATACACACACACAGATAGGCTTCAGATCAAC 2279
Qy 2521 ATGCATCACATCAAGTGTAAATCTTCAAAGTGTCTTCTTCTTCTCTCTCTTTTAT 2580
Db 2280 ATGCATCACATCAAGTGTAAATCTTCAAAGTGTCTTCTTCTTCTCTCTTTTAT 2339
Qy 2581 TGTGTGCTTTTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2640
Db 2340 TGTGTGCTTTTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2399
Qy 2641 GCTAGAGCTAAATCATATAGAAATGATGTATCTTGTGTGTGAGAAAGCCAGCT 2700
Db 2400 GCTAGAGCTAAATCATATAGAAATGATGTATCTTGTGTGTGAGAAAGCCAGCT 2459
Qy 2701 GGCTTAAGTTCACATTTTGTCCAGTGGCTAGCTTCCACCCAGCAGCTCCCAAAAT 2760
Db 2460 GGCTTAAGTTCACATTTTGTCCAGTGGCTAGCTTCCACCCAGCAGCTCCCAAAAT 2519
Qy 2761 GAAAGACCACTGTCAAGCAGCAGTCAAGAGTGTATGTGTGTGTGAGAAAGCCAGCT 2820
Db 2520 GAAAGACCACTGTCAAGCAGCAGTCAAGAGTGTATGTGTGTGTGAGAAAGCCAGCT 2579
Qy 2821 CCATCATTTGCTTCCCTCTCCACCCCTGTGACCTTACCTTAATCCCATTTGGGAG 2880
Db 2580 CCATCATTTGCTTCCCTCTCCACCCCTGTGACCTTACCTTAATCCCATTTGGGAG 2639
Qy 2881 CCAGGCAATGTTTGTCTTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db 2640 CCAGGCAATGTTTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2699
Qy 2941 CCAGTATGCTTCTTCTTCTTCAAGAGTGCATATGTTTCTTCTCAAAATAGTGT 3000
Db 2700 CCAGTATGCTTCTTCTTCTTCAAGAGTGCATATGTTTCTTCTCAAAATAGTGT 2759
Qy 3001 ATTCTTAAATTAACCAAAAGACAGGCAATCCACAGTGTGTGAGCATGAATCAAGCTTC 3060
Db 2760 ATTCTTAAATTAACCAAAAGACAGGCAATCCACAGTGTGTGAGCATGAATCAAGCTTC 2819
Qy 3061 ATTGTGTGAGTGTGAGTGGGATAAAGTGGATGTGTGAGAGAGTGGAAATCAAACTC 3120
Db 2820 ATTGTGTGAGTGTGAGTGGGATAAAGTGGATGTGTGAGAGAGTGGAAATCAAACTC 2879

Db	2326	GAGGAAAGGCCAGCTGSCCTAAAGTTCCACACTTTGTCCAGAGTGGCCCTAGACTCCACCCA	2388
Qy	2746	GCACAGCTCCCAAAATGAAAGACCACTCTCAAGCAGCAGTCAAGAGTGTGATGTCCACCC	2805
Db	2386	GCACAGCTCCCAAAATGAAAGACCACTCTCAAGCAGCAGTCAAGAGTGTGATGTCCACCC	2445
Qy	2826	ATCAGTATTTTTTTCATCATGTGCTGTGCTCTGCTCCCTTCCACACCCGTGTGACGT	2865
Db	2446	ATCAGTATTTTTTTCATCATGTGCTGTGCTCTGCTCCCTTCCACACCCGTGTGACGT	2505
Qy	2866	AATCGCATTTGGGAAGCCAGGACAATGTTTGGTGTCTGCTTTGGGTAAAGGAGCTCCCTG	2925
Db	2506	AATCGCATTTGGGAAGCCAGGACAATGTTTGGTGTCTGCTTTGGGTAAAGGAGCTCCCTG	2565
Qy	2926	AAGCTGTGGGCTCTCCATATGCTATGCTGTGCTCTGCTCCCTTCCACACATGCTTCTT	2985
Db	2566	AAGCTGTGGGCTCTCCATATGCTATGCTGTGCTCTGCTCCCTTCCACACATGCTTCTT	2625
Qy	2986	CAGAAACACATATGATCTCTTAAATACCCAGACAGGAGTCCACAGTGTGTGAGCA	3045
Db	2626	CAGAAACACATATGATCTCTTAAATACCCAGACAGGAGTCCACAGTGTGTGAGCA	2685
Qy	3046	TGAATCACAGCCTGCAATGTGTGAGTGTGAATAGTGGGATAAAAGTGGATGTCCAGAAAG	3105
Db	2686	TGAATCACAGCCTGCAATGTGTGAGTGTGAATAGTGGGATAAAAGTGGATGTCCAGAAAG	2745
Qy	3106	TGGAAATCAACCTCTGCAAGCAATCTTTCTCTTCTCTGGAAGTGTNTAAGCAATACC	3165
Db	2746	TGGAAATCAACCTCTGCAAGCAATCTTTCTCTTCTCTGGAAGTGTNTAAGCAATACC	2805
Qy	3166	TGAAGTCTGTGTGTGTGGTGAACCCAGACTGTCAATCATATAAGACCCAGACTGTCAAT	3225
Db	2806	TGAAGTCTGTGTGTGTGGTGAACCCAGACTGTCAATCATATAAGACCCAGACTGTCAAT	2865
Qy	3226	GAAAAAANAAAAAANAAAAAANAAAAA 3255	
Db	2866	GAAAAAANAAAAAANAAAAAANAAAAA 2895	
RESULT 14			
AAFS5252			
ID	AAFS5252 standard; cDNA; 2918 BP.		
XX			
AC	AAFS5252.		
XX			
DT	25-MAY-2001 (first entry;		
XX			
DE	Nucleotide sequence of murine parkin2 with a frameshift mutation.		
XX			
KK	Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease;		
XX	Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;		
XX	Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;		
KK	Brain tumour; head trauma; stroke; vascular irregularity;		
XX	metabolic irregularity; ss.		
XX			
QS	Xus sp.		
XX			
EH	Location/Qualifiers		
Key	129..630		
CDS	/*tag= a		
FT	/*product= "truncated parkin2"		
FT			
XX			
PN	EP1081225-A..		
XX			
PD	07-MAR-2001.		
XX			
XX	30-AUG-1999; 99EP-0116766.		
XX			
PR	30-AUG-1999; 99EP-0116766.		
XX			
PA	(BIOF-) BIOFRONTIERA PHARY GMBH.		
XX			
PI	Luebbert H;		

XX WPI: 2001-212797/22.
CR P-PSDB; AAB67525.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
PT producing a transgenic non-human animal as an animal model for
PT neurodegenerative diseases
XX
PS Claim 3; Page 28-29; 62pp; English.
XX
CC The present sequence encodes a murine parkin2 polypeptide. The
CC polynucleotide sequence contains a frameshift mutation, leading to a
CC truncated protein. Mutations or deletions in the parkin2 gene cause
CC Parkinson's disease in humans. The human parkin2 gene is located in
CC gene region 9q25.2-27. Parkin2 polypeptides and polynucleotides are
CC useful for analysing neurodegenerative diseases. They are also useful
CC for testing the efficacy of the treatment of a neurodegenerative disease
CC such as Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes inducing
CC Parkinson's syndrome like toxins, drugs, brain tumours, head trauma,
CC stroke, vascular irregularities or metabolic irregularities, associated
CC with a less active or non-active parkin protein.
XX
SQ Sequence 2918 BP; 777 A; 724 C; 727 G; 690 T; 0 other;
Query Match 79.0%; Score 2571; D5 22; Length 2918;
Best Local Similarity 89.6%; Pred No 0;
Matches 2918; Conservative 0; Mismatches 0; Indels 337; Gaps 1;
1 CTCACGAGGGGAAAGGGGAGGAGGCGCTGGATGACTAAACCTCAGCAAAACGCTGGTGGG 60
D5 1 CTCACGAGGGGAAAGGGGAGGAGGCGCTGGATGACTAAACCTCAGCAAAACGCTGGTGGG 60
61 AGGCTCGGGCGGGCCAGTCCCGGCTAGCTCTCTCAGCGGAGCCACACCGCC 120
D5 61 AGGCTCGGGCGGGCCAGTCCCGGCTAGCTCTCTCAGCGGAGCCACACCGCC 120
62 AGGCTCGGGCGGGCCAGTCCCGGCTAGCTCTCTCAGCGGAGCCACACCGCC 120
D5 62 AGGCTCGGGCGGGCCAGTCCCGGCTAGCTCTCTCAGCGGAGCCACACCGCC 120
121 CGGTGACCATGATAGTGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGG 180
D5 121 CGGTGACCATGATAGTGTGTGTCAGGTTCAACTCCAGCTATGCTTCCAGTGGAGTGG 180
181 ATCTGACACAGCATCTTGAGCTCAAGGAAGTGTGTGTCAGGACAGGGGTTCCAG 240
D5 181 ATCTGACACAGCATCTTGAGCTCAAGGAAGTGTGTGTCAGGACAGGGGTTCCAG 240
182 ATCTGACACAGCATCTTGAGCTCAAGGAAGTGTGTGTCAGGACAGGGGTTCCAG 240
D5 182 ATCTGACACAGCATCTTGAGCTCAAGGAAGTGTGTGTCAGGACAGGGGTTCCAG 240
241 CTGACAGCTGGTGTGATTTTTCGCGGAGGAGGCTTCGAATCAGCTCAGCGTTCAAA 300
D5 241 CTGACAGCTGGTGTGATTTTTCGCGGAGGAGGCTTCGAATCAGCTCAGCGTTCAAA 300
242 CTGACAGCTGGTGTGATTTTTCGCGGAGGAGGCTTCGAATCAGCTCAGCGTTCAAA 300
D5 242 CTGACAGCTGGTGTGATTTTTCGCGGAGGAGGCTTCGAATCAGCTCAGCGTTCAAA 300
301 ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTC 360
D5 301 ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTC 360
302 ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTC 360
D5 302 ACTGTGACCTGGAAACACAGAGTATTGTACACATAGTACAGACACCGAGGAGAGTC 360
361 ATGAACAAATGCATCTCGAGGGGACGAAACCCAGAGACCTCAGAGGGCTCATATGGG 420
D5 361 ATGAACAAATGCATCTCGAGGGGACGAAACCCAGAGACCTCAGAGGGCTCATATGGG 420
362 ATGAACAAATGCATCTCGAGGGGACGAAACCCAGAGACCTCAGAGGGCTCATATGGG 420
D5 362 ATGAACAAATGCATCTCGAGGGGACGAAACCCAGAGACCTCAGAGGGCTCATATGGG 420
421 AGTCAGAGGCTTGACACAGTGGACCTGAGACGCCATACCTCCGGTGGAGCTGTGG 480
D5 421 AGTCAGAGGCTTGACACAGTGGACCTGAGACGCCATACCTCCGGTGGAGCTGTGG 480
422 AGTCAGAGGCTTGACACAGTGGACCTGAGACGCCATACCTCCGGTGGAGCTGTGG 480
D5 422 AGTCAGAGGCTTGACACAGTGGACCTGAGACGCCATACCTCCGGTGGAGCTGTGG 480
481 GGCTGGCGGTCAATCTGACACACAGAGTATAGAGGATTCAGAGCAGCCAGAGTCCAG 540
D5 481 GGCTGGCGGTCAATCTGACACACAGAGTATAGAGGATTCAGAGCAGCCAGAGTCCAG 540
482 GGCTGGCGGTCAATCTGACACACAGAGTATAGAGGATTCAGAGCAGCCAGAGTCCAG 540
D5 482 GGCTGGCGGTCAATCTGACACACAGAGTATAGAGGATTCAGAGCAGCCAGAGTCCAG 540
541 TTAACCCACCTACAAACAGCTTTTTCATCTACTGCAAGGCCCCCTGCCAAAGTCCAGC 600
D5 541 TTAACCCACCTACAAACAGCTTTTTCATCTACTGCAAGGCCCCCTGCCAAAGTCCAGC 600
542 TTAACCCACCTACAAACAGCTTTTTCATCTACTGCAAGGCCCCCTGCCAAAGTCCAGC 600
D5 542 TTAACCCACCTACAAACAGCTTTTTCATCTACTGCAAGGCCCCCTGCCAAAGTCCAGC 600
601 CTGGAAGCTCGAGTTTCAGTGTGGCAGCTCTGCAAAACAGCAACCTCAGCTTGGCCAGG 660
D5 601 CTGGAAGCTCGAGTTTCAGTGTGGCAGCTCTGCAAAACAGCAACCTCAGCTTGGCCAGG 660
602 CTGGAAGCTCGAGTTTCAGTGTGGCAGCTCTGCAAAACAGCAACCTCAGCTTGGCCAGG 660
D5 602 CTGGAAGCTCGAGTTTCAGTGTGGCAGCTCTGCAAAACAGCAACCTCAGCTTGGCCAGG 660

QY 561 GCCCATCTTCTGGGAGCA-TGTATTATTCGAAACCCGATGAGTGTGAGTGGCCATCTC 720
D5 561 ----- 657
QY 721 CAGACTGCCCTGGAAACGAGAGCTGAA-TTTTCTTTAAATGTGGAGCAGACCCACCTCAG 780
D5 721 ----- 657
QY 781 ACAAGGACACGCTGGTAGCTTTTGAACCTGATCAGCAGCAACAGCGCCGACATCCCTTGA 840
D5 781 ----- 657
QY 841 TAGCGTCACAGATGTCAGSAGCCCTGCTCTGTCTTCCAGGTGTAACCCGCTCAGCTGA 900
D5 841 ----- 657
QY 901 TCTGTTGGACTGTTTCCACTTGTATTTGTCACAAAGCTCAACAGATCGSCAGTTGTCC 960
D5 901 ----- 657
QY 961 ACGATGCTCAACTTGGCTACTTCTCTCCGCTGTGTAGCTGCTCTCCAACTCTCCTATTA 1020
D5 961 ----- 683
QY 1021 AAGAGCTCCATCACTTCAGGATCCTTGGAGAGCAGTACACTAGTACCGAGCAGTATG 1080
D5 984 AAGAGTCCATCACTTCAGGATCCTTGGAGAGCAGTACACTAGTACCGAGCAGTATG 743
QY 1081 GGGCCGAGGAATCGTCTGCAAAATGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
D5 744 GGGCCGAGGAATCGTCTGCAAAATGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 803
QY 1141 CTGAGCTGCTACTGACAGGCGCAGAGGAAATCAGCTGGGAGAGGGGAGGAGGAGGAGG 1200
D5 924 CTGAGCTGCTACTGACAGGCGCAGAGGAAATCAGCTGGGAGAGGGGAGGAGGAGGAGG 863
QY 1201 CTTGCGGCTTGT 1260
D5 864 CTTGCGGCTTGT 923
QY 1261 CACTGCTGCAACCTCAGSAGCCACTTCTCAGCTTACAGGTTGAGAGGAGGAGGAGGAGG 1320
D5 924 CACTGCTGCAACCTCAGSAGCCACTTCTCAGGCTTACAGGTTGAGAGGAGGAGGAGGAGG 983
QY 1321 AGCAAGCTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
D5 984 AGCAAGCTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1043
QY 1381 CTGCTGCAACCTCAGSAGCCACTTCTCAGCTTACAGGTTGAGAGGAGGAGGAGGAGGAG 1440
D5 1244 CTGCTGCAACCTCAGSAGCCACTTCTCAGGCTTACAGGTTGAGAGGAGGAGGAGGAGGAG 1103
QY 1441 CCGAGTGCAGCTGGAGTGGT 1500
D5 1244 CCGAGTGCAGCTGGAGTGGT 1163
QY 1501 GAGATCAGCTGTTGAGCTGTAGAGAGAGTGTCACTTGGCCCTGAGAGGAGGAGGAGGAG 1560
D5 1164 GAGATCAGCTGTTGAGCTGTAGAGAGAGTGTCACTTGGCCCTGAGAGGAGGAGGAGGAG 1223
QY 1561 GGGAAATCCGAGAGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
D5 1224 GGGAAATCCGAGAGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1283
QY 1621 CATTAATTCGACACACAAACACAGGCTGCAGATTACAGAAACAGGAGGAGGAGGAGGAG 1680
D5 1284 CATTAATTCGACACACAAACACAGGCTGCAGATTACAGAAACAGGAGGAGGAGGAGGAG 1343
QY 1691 TAGGGACCCACAGAAACACAGAGCAGGCTGGCCCGAGGGGAGGAGGAGGAGGAGGAGG 1740
D5 1344 TAGGGACCCACAGAAACACAGAGCAGGCTGGCCCGAGGGGAGGAGGAGGAGGAGGAGG 1403

[illegible]

Sequence 93, Application US/08505486
 Patent No. 5955573
 GENERAL INFORMATION:
 APPLICANT: Jesse M. Jaynes
 TITLE OF INVENTION: UBICUTIN-LYTIC PEPTIDE FUSION GENE
 TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
 TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
 NUMBER OF SEQUENCES: 98

ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
 CORRESPONDENCE ADDRESS:
 STREET: 555 Thirteenth Street N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: D-SKETTE, 3.5 INCH, 1.4 MB STORAGE
 COMPUTER: IBM COMPATIBLE
 OPERATING SYSTEM: DOS
 SOFTWARE: WordPerfect 5.1+

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/505,486
 FILING DATE: 21-JUL-1995
 CLASSIFICATION: 536

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: U.S. 08/279,472
 FILING DATE: 22-JUL-1994
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: WALKER, BARBARA W.
 REGISTRATION NUMBER: 35,400
 REFERENCE/DOCKET NUMBER: 2093-117A
 TELEPHONE: (202) 783-6040
 TELEFAX: (202) 783-6031

INFORMATION FOR SEQ ID NO: 93:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1154
 TYPE: NUCLEIC ACID
 STRANDEDNESS: DOUBLE STRANDED
 TOPOLOGY: LINEAR

MOLECULE TYPE:
 DESCRIPTION: GENOMIC DNA
 S-28-505-486-93

Query Match 1.7%; Score 56.4; DB 2; Length 1154;
 Best Local Similarity 52.6%; Pred. No. 1e-05;
 Matches 123; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

116 CGCCCGGTGACCATGATGTTTCTAGTTCACCTCCAGCTATGCTTCCCGAGTGA 175
 908 CTCCAGGCGAGATCGATCTCTGTGAGACCTTAAGGGGAGAGAGATCACCCCTAGA 957
 176 GGTGATTTCTGACACGAGCATTTTGGAGCTCAAGGAAGTGGTTGCTAAGCGACAGGGGGT 235
 968 GGTGATCTTCGACACCATCGACATGTCGAAGCCAGATCCAGGACAGGAGGAT 1027
 236 TCCAGCTGACAGCTGGGTGATTTTCCCGGGAGAGAGCTTCGATACCTGAGGT 295
 1028 TCCCGACAGCAGCGGTTTGAATTTCCCGGAAAGAGCTTGAGGATGTTGCTACTCT 1087
 296 TCAAACTGTGACCTGGAAACAAGAGATTTTACACATAGTACAGAGACCAAG 349
 1088 TGCGGACTACACATCCAGAGGAGTCAATCTCCATCTCGTGCTCCGTCTCCG 1141

RESULT 10
 -08-801-028-93
 Sequence 93, Application US/08801028
 Patent No. 6018102
 GENERAL INFORMATION:
 APPLICANT: JOAN GARBARINO
 APPLICANT: JESSE M. JAYNES

APPLICANT: WILLIAM BELKNAP
 TITLE OF INVENTION: UBICUTIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCT
 NUMBER OF SEQUENCES: 98
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STEVEN J. HULTQUIST
 ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
 STREET: 200 PARK DRIVE, SUITE 210
 STREET: P.O. BOX 14329
 CITY: RESEARCH TRIANGLE PARK
 STATE: NORTH CAROLINA
 COUNTRY: USA
 ZIP: 27709

COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
 COMPUTER: APPLE MACINTOSH
 OPERATING SYSTEM: MACINTOSH
 SOFTWARE: V.S. WORD 5.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/801,028
 FILING DATE: 19-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/279,472

FILING DATE: JULY 22, 1994
 APPLICATION NUMBER: 08/225,476
 FILING DATE: 04-23-94
 APPLICATION NUMBER: 08/225,476
 FILING DATE: 04-08-94
 APPLICATION NUMBER: 08/039,620

FILING DATE: 06-04-93
 APPLICATION NUMBER: 08/149,491
 FILING DATE: 11-08-93
 APPLICATION NUMBER: 08/149,889
 FILING DATE: 11-08-93

ATTORNEY/AGENT INFORMATION:
 NAME: WASSERMAN, PRAN S.
 REGISTRATION NUMBER: 34,273
 REFERENCE/DOCKET NUMBER: 4013-104
 TELEPHONE: (919) 990-9531
 TELEFAX: (919) 990-9532

INFORMATION FOR SEQ ID NO: 93:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1154
 TYPE: NUCLEIC ACID
 STRANDEDNESS: DOUBLE STRANDED
 TOPOLOGY: LINEAR

MOLECULE TYPE:
 DESCRIPTION: GENOMIC DNA
 US-08-801-028-93

Query Match 1.7%; Score 56.4; DB 2; Length 1154;
 Best Local Similarity 52.6%; Pred. No. 1e-05;
 Matches 123; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

116 CGCCCGGTGACCATGATGTTTGTGAGTTCACCTCCAGCTATGCTTCCCGAGTGA 175
 908 CTCCAGGCGAGATCGATCTCTGTGAGACCTTAAGGGGAGAGAGATCACCCCTAGA 957
 176 GGTGATTTCTGACACGAGCATTTTGGAGCTCAAGGAAGTGGTTGCTAAGCGACAGGGGT 235
 968 GGTGATCTTCGACACCATCGACATGTCGAAGCCAGATCCAGGACAGGAGGAT 1027
 236 TCCAGCTGACAGCTGGGTGATTTTCCCGGGAGAGAGCTTCGATACCTGAGGT 295
 1028 TCCCGACAGCAGCGGTTTGAATTTCCCGGAAAGAGCTTGAGGATGTTGCTACTCT 1087
 296 TCAAACTGTGACCTGGAAACAAGAGATTTTACACATAGTACAGAGACCAAG 349
 1088 TGCGGACTACACATCCAGAGGAGTCAATCTCCATCTCGTGCTCCGTCTCCG 1141

RESULT 1:
 US-09-340-154-93

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX nucleic - nucleic search, using sw model

Run on: October 24, 2003, 12:19:43 ; Search time 599 Seconds
(without alignments)
14573.058 Million cell updates/sec

Title: US-09-830-703-1

Perfect score: 3255

Sequence: 1 ctccagcagggaggaggga.....aaaaaaaaaaaaaaaaaaaa 3255

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1792395 seqs, 134090451 residues

Total number of hits satisfying chosen parameters: 3584730

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	340.2	10.5	471	10	US-09-785-548-3
2	109.8	3.4	449	9	US-03-864-761-3503
3	103.2	3.2	140	9	US-09-864-761-3503
4	65.4	2.0	810	12	US-09-814-353-5846
5	65.4	2.0	810	12	US-09-814-353-12127
6	65	2.0	65	12	US-09-908-975-26543
7	62.8	1.9	143601	12	US-09-855-824-3
8	60.4	1.9	508	12	US-09-814-353-18511
9	59.8	1.8	440	11	US-09-918-995-35452
10	59.8	1.8	473	11	US-09-918-995-23749
11	59.8	1.8	505	11	US-09-918-995-20526
12	59.8	1.8	1285	10	US-09-997-107-45
13	59.2	1.8	1734	11	US-09-946-374-51
14	59.2	1.8	1734	12	US-10-015-387A-51
15	59.2	1.8	1734	12	US-10-063-735-51
16	59.2	1.8	1734	12	US-10-006-130A-51

C 17	59.2	1.8	1734	12	US-10-199-372-201	Sequence 201, App
C 18	59.2	1.8	1734	12	US-10-216-163-159	Sequence 155, App
C 19	59.2	1.8	1734	12	US-10-226-172A-51	Sequence 51, App1
C 20	59.2	1.8	1734	12	US-10-187-749-201	Sequence 201, App
C 21	59.2	1.8	1734	12	US-10-194-457-201	Sequence 201, App
C 22	59.2	1.8	1734	12	US-10-154-642-201	Sequence 201, App
C 23	59.2	1.8	1734	12	US-10-136-747-201	Sequence 201, App
C 24	59.2	1.8	1734	12	US-10-015-382A-51	Sequence 51, App1
C 25	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 26	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 27	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 28	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 29	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 30	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 31	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 32	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 33	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 34	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 35	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 36	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 37	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 38	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 39	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 40	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 41	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 42	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 43	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 44	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App
C 45	59.2	1.8	1734	12	US-10-173-889-201	Sequence 201, App

ALIGNMENTS

RESULT 1
US-09-785-548-3
; Sequence 3, Application US/09785548
; Patent No. US2002015557A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF P
; REFERENCE: STG0005
; CURRENT APPLICATION NUMBER: US/09/785,548
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(471)
US-09-785-548-3

Query Match	10.5%	Score	340.2	DB 10	Length	471			
Best Local Similarity	84.3%	Pred	NC. 5.6e-86						
Mismatches	334	Conservative	0	Mismatches	73	Indels	0	Gaps	0
QV	538	CAGTTAAACCCACCTACACAGCCTTTTCATCTCTCTGCAAGGCCCTGCGACAGGTCC	597						
DB	11	CAGGTAGATCAATCTACACAGCCTTTTATGTATTGCAAGGCCCTGTCAAGAGTGC	70						
QV	598	AGCTCGAAGTCCGAGTTCAGTGTGGACCTCCAAACAGCAACCTCAGCTTGGCCC	657						
DB	71	AGCGGGAAATCTAGGTACAGTCAGCAGCCTGCGACGCGCAACCTCAGCTTGGCCC	130						
QV	658	AGCGCCCATCTTCCTGGACAGATGCTTTTAATTCCAAACCGATGAGTGGTGGCCAGT	717						
DB	131	AGGTCCTATCTTGCCTGGATGATGTTTAAATCCAAACCGGATGAGTGGTGAATGCCAAT	190						
QV	718	CTCCAGACTGCCCTGGACACAGAGCTGAATTTTCTTTAAATGTGGAGCAGACCCCAACT	777						

```

Db 191 CCCCACACTGCCCTGGAGTAGTCAGAAATTTCTTTAAATGTGGAGCACACCCCACTT 250
QY 778 CACAGAGACACCTCGGTAGCTTTTGAACCTGATCACAGCAACAGCGGAGCATCCCTT 537
Db 251 CTGACANGAACATCAGTAGCTTGCACCTGATCGCAAAATAGTCGGAACTACTT 310
QY 838 GCATAGGTGCACAGATGTACAGAGCCCTGCTCTGGTCTTTCAGTGTAAACACCGTCAG 897
Db 311 GCATTAGTGCACAGACGTTCAGAGCCCGTCTCTGGTCTTTCAGTGTAACTCCCGCCACG 370
QY 898 TGATCTGTTGGACTGTTCCACTTGTATGTGTACAAAGACTCAACGATCGGCAGTTTG 957
Db 372 TGATTTCCTAGACTGTTCCACTTATAGTGTGACAAAGACTCAATGATGGCAGTTTG 430
QY 958 TCACGATGTCAACTTGCTACTCTCCCTGCCGTGT 994
Db 431 TTCACGACCTCAACTGGCTACTCTCCCTGCCGTGT 467

RESULT 2
US-09-864-761-3503/c
; Sequence 3503, Application: US/09864761
; Patent No. US20020348763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3503
; LENGTH 449

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALC35697.18
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.63
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
US-09-864-761-3503

Query Match 3.44; Score 109.8; DB 9; Length 449;
Best Local Similarity 84.84; Pred. No. 2.2e-20;
Matches 123; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1066 GGTACACAGCACTGTGGGCGGAGGATGCTGCTGCAAAATGGGAGGTGTGCTGTGCCCC 1:25
Db 445 GGTACACAGCACTGTGGGCGGAGGATGCTGCTGCAAAATGGGAGGTGTGCTGTGCCCC 396
QY 1126 GTCTTGGCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1:185
Db 385 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326
QY 1196 GGGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
Db 325 GGGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

RESULT 3
US-09-864-761-20274/c
; Sequence 20274, Application: US/09864761
; Patent No. US20020348763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3503
; LENGTH 449

```



```

: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: US 60/211,940
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: US 60/216,820
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US 60/220,661
: PRIOR FILING DATE: 2000-07-15
: PRIOR APPLICATION NUMBER: US 60/257,672
: PRIOR FILING DATE: 2000-12-21
: NUMBER OF SEQ IDS NOS: 22037
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 18511
: LENGTH: 508
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-814-353-185.11

```

Query Match 1.84; Score 59.8; DB 11; Length 505;
 Best Local Similarity 53.18; Pred. No. 4e-06;
 Matches 127; Conservative 0; Mismatches 112; Indels 0; Gaps 2;
 QY 126 ACCATGATAGTGTTCAGCTTCACTCCAGCTATGGTTCACAGTGGAGTGATCT 185
 DB 55 AACATGCGAGATCTTTGAGACCCCTACTGGCAAAACCATCACCTTGAGTGGAGCCC 114
 QY 186 GACACCCAGCTTTCAGCTCAGGAAGTGTTCAGGACAGGGGTTCAGCTGAC 245
 DB 113 AGTGACACCATGTGAGATGTCAAGCCAAAATTCAAGCAAGGAGGTATCCACCTGAC 174
 QY 246 CAGCTGCGGTGTGATTTTCGCGGAAGGAGCTTCGCAATCACTGACGCTTCAAACTGT 305
 DB 175 CAGGACGCTGTGATTTTCGCGGAAGGAGCTTCGCAATCACTGACGCTTCAAACTGT 234
 QY 306 GACCTGGACACACAGAGTATTGTACACATAGTACAGACACCGGAGAGAGTATGA 364
 DB 235 AACATCCAGAAAGAGTCCACCTGACCTGGTTCGCGCTCGAGGTGGCATATTGA 293

RESULT 12
 US-09-987-107-45
 Sequence 45, Application US/09987107
 Patent No. US20020156007A1
 GENERAL INFORMATION:
 APPLICANT: GRAVERSEN, Jonas
 APPLICANT: MOESTRUP, Soren
 TITLE OF INVENTION: APOLOPOROTEIN ANALOGUES
 FILE REFERENCE: GRAVERSENIA
 CURRENT APPLICATION NUMBER: US/09/987,107
 PRIOR FILING DATE: 2001-11-13
 PRIOR APPLICATION NUMBER: US 60/264,022
 PRIOR FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: DK PA2001 00057
 PRIOR FILING DATE: 2001-11-15
 PRIOR APPLICATION NUMBER: DK PA2000 01682
 PRIOR FILING DATE: 2000-11-10
 NUMBER OF SEQ ID NOS: 91
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 45
 LENGTH: 1285
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (100)..(1110)
 OTHER INFORMATION:
 US-09-987-107-45

Query Match 1.84; Score 59.8; DB 10; Length 1285;
 Best Local Similarity 51.38; Pred. No. 7.6e-06;
 Matches 139; Conservative 0; Mismatches 132; Indels 0; Gaps 5;
 QY 136 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 195
 DB 137 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 195
 QY 196 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 255
 DB 197 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 255
 QY 256 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 315
 DB 257 TGTGTTCAGGTTCAACTCCAGCTATGGCTTCAGTGGAGGTTCGATTCGACACCA 315
 QY 316 AACAGATTTGTACACATAGTACAGACACCGAGGAGTGTGAAACAAATGAT 375
 DB 317 AGAGTCTACTTCTATCTTGTGTGAGACTTCGTGTGTGATCCATCGAGGTAGGGGT 375
 QY 376 CTGAGGGGACCAACCCCGAGACCTCAGA 406

DB 377 GATGTGATGAACCCCGCCAGAGCCCTGGGA 407
 RESULT 13
 US-09-946-374-51/c
 Sequence 51, Application US/09946374
 Publication No. US2003007322A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grima, Di, Christopher C.
 APPLICANT: Gurley, Austin L.
 APPLICANT: Hillan, Kenneth C.
 APPLICANT: Pan, James
 APPLICANT: Pacini, Nicholas P.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Turas, Danie.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William J.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: 2283261C1
 CURRENT APPLICATION NUMBER: US/09/946,374
 CURRENT FILING DATE: 2001-09-04
 PRIOR APPLICATION NUMBER: 60/098716
 PRIOR FILING DATE: 1998-09-01
 PRIOR APPLICATION NUMBER: 60/098723
 PRIOR FILING DATE: 1998-09-01
 PRIOR APPLICATION NUMBER: 60/098749
 PRIOR FILING DATE: 1998-09-01
 PRIOR APPLICATION NUMBER: 60/098750
 PRIOR FILING DATE: 1998-09-01
 PRIOR APPLICATION NUMBER: 60/098903
 PRIOR FILING DATE: 1998-09-02
 PRIOR APPLICATION NUMBER: 60/098921
 PRIOR FILING DATE: 1998-09-02
 PRIOR APPLICATION NUMBER: 60/098943
 PRIOR FILING DATE: 1998-09-02
 PRIOR APPLICATION NUMBER: 60/099516
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099556
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099598
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099602
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099642
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099741
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099754
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099763
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099792
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099808
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099812
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099815
 PRIOR FILING DATE: 1998-09-10

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: October 24, 2003, 09:39:33 ; Search time 4317 Seconds

(without alignments)
 18325.474 Million cell updates/sec

Title: US-09-830-703-1

Perfect score: 3255

Sequence: - cccagcgaggggaaggggga.....aaaaaaaaaaaaaaaaaaaa 3255

Scoring table: IDENTITY NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2278132 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 4562754

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:

1: em_estba:

2: em_estum:

3: em_estin:

4: em_estm:

5: em_estov:

6: em_estpl:

7: em_estro:

8: em_hci:

9: gb_est1:

10: gb_est2:

11: gb_hci:

12: gb_est3:

13: gb_est4:

14: gb_est5:

15: em_estfun:

16: em_estcr:

17: em_gss_hum:

18: em_gss_inv:

19: em_gss_pin:

20: em_gss_ttt:

21: em_gss_fun:

22: em_gss_mam:

23: em_gss_mus:

24: em_gss_pro:

25: em_gss_rod:

26: em_gss_pbg:

27: em_gss_vri:

28: gb_gss1:

29: gb_gss2:

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1929.8	58.7	2182	11	AK015990
2	590.8	18.2	963	14	BY715565
3	459.4	14.1	913	10	BG720953
4	438.8	13.5	453	13	BY596928

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	434.4	13.3	434	13	BY592998
6	331.8	12.0	445	10	BH848509
7	388.8	11.9	393	13	BY232905
8	358	11.0	419	13	BY237113
9	349.4	10.7	351	13	BY235535
10	343.4	10.2	404	14	CB811663
11	302	9.3	142	14	CD495236
12	231.8	9.0	415	12	BM538799
13	282.4	8.7	400	3	A1228346
14	273.8	8.4	391	10	S285222
15	272.4	8.4	274	10	BE543338
16	271.6	8.3	605	9	A1214045
17	264.2	8.1	642	12	SV487248
18	255.8	7.9	455	14	CB741196
19	232.2	7.7	566	10	BG573071
20	238.5	7.6	443	14	CE789229
21	217.4	7.6	332	10	BE235492
22	245	7.5	289	9	AV268904
23	235.8	7.3	536	13	HJ380821
24	236.9	7.3	804	13	BU249355
25	232.4	7.1	358	10	BF542359
26	228.4	7.0	295	10	BF524700
27	218.6	6.2	1142	14	CD498235
28	213	6.5	488	10	BF095769
29	228.8	6.4	481	9	AA875183
30	228.8	6.4	597	14	CB613373
31	228.8	6.4	600	14	CB583527
32	198.2	6.1	345	13	BQ184748
33	197.2	6.1	710	13	BQ186216
34	181.8	5.6	279	9	AV237665
35	181.5	5.6	214	9	A1211704
36	166.4	5.5	400	14	CB699223
37	170.2	5.3	873	14	CB512427
38	166.4	5.1	227	10	BF594017
39	155	4.8	363	10	BH870131
40	154.2	4.7	531	10	BF548754
41	143.8	4.3	973	10	BF788569
42	131.4	4.0	499	10	BF550135
43	114	3.5	606	10	BG617830
44	112	3.4	719	28	AG357591
45	126.8	3.3	648	9	AU177228

ALIGNMENTS

RESULT 1
 AK015990
 LOCUS
 DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1910S38K15 product:ipafkin, full insert sequence.
 ACCESSION AK015990
 VERSION AK015990.1 GI:12854554
 KEYWORDS RTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

2182 bp -RNA linear HTC 05-DEC-2002
 Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1910S38K15 product:ipafkin, full insert sequence.

Gennanci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999);
 Genom. Res. 10 (10), 1617-1630 (2000);
 Gennanci, P., Shibata, Y., Hayatsu, M., Sugahara, Y., Shibata, K.,
 Koh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000);
 PUBMED 11042159
 REFERENCE 3

1830	QY	CAAGGGTTGTTGGCCNCCGCTTGGGTAAAGTAAACCTTCTGAGGATTTCTCAGAAAGGACACACTT	1889
785	DB	CCAGGGTGTGTGCACGCGCTTGGGTAAAGTAAACCTTCTGAGGATTTCTCAGAAAGGACACACTT	944
1890	QY	GAGATCTCAGGAAAGCGTCTCTCATGTGTAATATCATCTATTCCTCCAAAGGCGCCCTCTCAG	1949
845	DB	GAGATCTCAGGAAAGCGTCTCTCATGTGTAATATCATCTATTCCTCCAAAGGCGCCCTCTCAG	904
1950	QY	TAGTCAAAACTATTGTGTTATCCCTCCCAATCTCTATTTACAAATGCTGCTGATGAGAT	2009
905	DB	TAGTCAAAACTATTGTGTTATCCCTCCCAATCTCTATTTACAAATGCTGCTGATGAGAT	964
2010	QY	TACACCCCTCTGTGTACTAATCAGCTTATACACCAAGTGTAGAACCTTAGGAAAGCTAACT	2069
965	DB	TACACCCCTCTGTGTACTAATCAGCTTATACACCAAGTGTAGAACCTTAGGAAAGCTAACT	1024
2070	QY	GGATGGCAGACTGCTTAAATCGCAGGAGGAGCTCAGAAAGCCAAACTACTTCCGTTCCGTT	2129
1025	DB	GGATGGCAGACTGCTTAAATCGCAGGAGGAGCTCAGAAAGCCAAACTACTTCCGTTCCGTT	1084
2130	QY	TCAATATCTGGAACCTTTAGAAAGAAATGATCTTTTTTCCCTCCAAAGATTAACAAGT	2189
1085	DB	TCAATATCTGGAACCTTTAGAAAGAAATGATCTTTTTTCCCTCCAAAGATTAACAAGT	1144
2190	QY	CTGCAATTTGGTGTGGAGTATTCCTATCTGCAGCTGTGAAAGTTTAGCTTCACTGTGAATTT	2249
1145	DB	CTGCAATTTGGTGTGGAGTATTCCTATCTGCAGCTGTGAAAGTTTAGCTTCACTGTGAATTT	1204
2250	QY	AACAGAGAAAGTGGCTATAAAGGGGGGCTTTTAAAGAGACAATCCCATGATCGTGGCCCA	2309
1205	DB	AACAGAGAAAGTGGCTATAAAGGGGGGCTTTTAAAGAGACAATCCCATGATCGTGGCCCA	1264
2310	QY	ATGCTAAACAAGGGGTCAAGAAACACAAATGTTTATAGAGAGGAGCATCCCTCGACGATCTG	2369
1265	DB	ATGCTAAACAAGGGGTCAAGAAACACAAATGTTTATAGAGAGGAGCATCCCTCGACGATCTG	1324
2370	QY	AATGAGAGTATGCCAGCCCTTCCACCAAGTGGGACACCTCTGATATATCTGCTGCC	2429
1325	DB	AATGAGAGTATGCCAGCCCTTCCACCAAGTGGGACACCTCTGATATATCTGCTGCC	1384
2430	QY	TCTCTGTGTTAAAGCCCGAGGAGCCCATCCACCCAGTGGTCTACAGACAGGCGCAAT	2489
1385	DB	TCTCTGTGTTAAAGCCCGAGGAGCCCATCCACCCAGTGGTCTACAGACAGGCGCAAT	1444
2490	QY	ACACACACACCAAGATAGCTTCTAGATCAACATGATCAACATCAAGTGTAAATCTCTCA	2549
1445	DB	ACACACACACCAAGATAGCTTCTAGATCAACATGATCAACATCAAGTGTAAATCTCTCA	1504
2550	QY	AGGTTTCTTTCTTTCTGTTTCTTTTATATTTGTTTGGCTTTGGC-TCTTTTCTTTCTT	2609
1505	DB	AGGTTTCTTTCTTTCTGTTTCTTTTATTTGTTTGGCTTTGGCTTTTCTTTCTTTCTT	1564
2608	QY	TTTTTGTGTGTGGGGCTACCAAACTTGAGGCGCTAGAGCTAAAAATCATATAGAAAT	2667
1565	DB	TTTTTGTGTGTGGGGCTACCAAACTTGAGGCGCTAGAGCTAAAAATCATATAGAAAT	1624
2668	QY	GATGTTATCTTGTGGTGTGAGGAAGCGCAGCTGGGCTAAGTGTCAACATTTTGTCCCAAT	2727
1625	DB	GATGTTATCTTGTGGTGTGAGGAAGCGCAGCTGGGCTAAGTGTCAACATTTTGTCCCAAT	1684
2728	QY	GGCCTTAGACTCCACCCAGCAGTCCCAAAATGAAAGACCACTGTCAAGCAGCAGCTC	2787
1685	DB	GGCCTTAGACTCCACCCAGCAGTCCCAAAATGAAAGACCACTGTCAAGCAGCAGCTC	1744
2788	QY	AGGAGTCTGATGTACCCCATCACTATTTTTTTTCCATCATGTGTGCTTGCCTCTGCCTCCT	2847
1745	DB	AGGAGTCTGATGTACCCCATCACTATTTTTTTTCCATCATGTGTGCTTGCCTCTGCCTCCT	1804
2848	QY	TCCACACCGGTGTGACGTAAATGCGATTGTGGGAAGCGAGGCAATGTTTGCATGTTCTGTTT	2907
1805	DB	TCCACACCGGTGTGACGTAAATGCGATTGTGGGAAGCGAGGCAATGTTTGCATGTTCTGTTT	1864
2908	QY	GGGTAAAGGAGCTCCCTCGAAGCTCTGTGGCTCTCCAGTATGGTCTGCTTTTTTTTCTCTAAC	2967

[illegible]

Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, S., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,
S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,
Ishii, Y., Itoh, M., Kagawa, T., Kawa, J., Kohma, Y., Kondo, S., Kom-
H., Koya, S., Miyazaki, A., Murata, M., Nakamura, N., Nomura, K.,
Nimazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sato, H.,
Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tegami, M., Takeda, Y.,
Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001).
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000).
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000).

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001).
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

```

source
1..963
  /organism="Mus musculus"
  /mol_type="cDNA"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="4930538X15"
  /sex="male"
  /tissue_type="testis"
  /dev_stage="adult"
  /ab_tost="SOLR"
  /clone_1b="RIKEN full-length enriched adult male testis"
  /notes="Site 1: XhoI; Site 2: BamHI. cDNA library was
  prepared and sequenced in Mouse Genome Encyclopedia
  Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in
  RIKEN. Division of Experimental Animal Research in Riken
  contributed to prepare mouse tissues. 1st strand cDNA was
  primed with a primer [5',
  GAGAGAGAGAGGATCAAGAGCTCTTTTCTTTTCTTTT 3']. cDNA was
  prepared by using trehalose thermo-activated reverse
  transcriptase and subsequently enriched for full-length by
  cap-trapper. Second strand cDNA was prepared with the
  primer adapter of sequence [5',
  GAGAGAGAGAGGCGCGCAATTAATCTCGAGTTAATTAATTAATCCCCCCCC
  3']. cDNA was cloned into the XhoI and BamHI sites."
BASE COUNT      246 a 249 c 255 g 212 t
ORIGIN
Query Match
Best Local Similarity 95.31; Pred. NC. 1.1e-47;
Matches 653; Conservative 0; Mismatches 27; Indels 6; Gaps 5;
QY 1230 CAGCCCTACAGGTGGACAAAGAGCGCTGTGACGAGCTCGTGGAGAGCGCTCCAAAG 1349
DB 245 CAGCGCTACAGGTGGACAAAGAGCGCTGTGACGAGCTCGTGGAGAGCGCTCCAAAG 304
QY 1350 GAAACCATCAGAGAGCCACAGCTTCCTCCCTGACAGCTCCCAATGAAAGAAC 1409
DB 305 GAAACCATCAGAGAGCCACAGCTTCCTCCCTGACAGCTCCCAATGAAAGAAC 364
QY 1410 CGAGGATGTATGCACATGAAGTGTCTTCAGCCCTCAGTCAAGCTGGAGTGTGTGGAAC 1469

```

```

DB 1365 CGAGGATGTATGCACATGAAGTGTCTTCAGCCCTCAGTCTCAAGCTGTGTGTGGAAC 424
QY 1470 TGTGGTGTGAGTGGAAACGAGCTGCTGATGGGAAATCATCTGTTTGAGCTGTAGAGAGAG 1529
DB 425 TGTGGCTGTGAGTGGAAACGAGCTGCTGATGGGAAATCATCTGTTTGAGCTGTAGAGAGAG 484
QY 1510 ATGTCACTTGGCTGGAGCGACACACCTTCAAGGAAATCTCCGAAAGATTCCTTACCTTCTT 1569
DB 485 ATGTCACTTGGCTGGAGCGACACACCTTCAAGGAAATCTCCGAAAGATTCCTTACCTTCTT 544
QY 1590 AGGCACTTCTTCTTCGATGATATATAGACATTAATTCGCGACACACACACACAGGCTG 1649
DB 545 AGGCACTTCTTCTTCGATGATATATAGACATTAATTCGCGACACACACACACAGGCTG 604
QY 1650 CAGATTACAGAGACACACCTTCTGATGCTTCTTCTAGGGGACCCACACAGACACACACAC 1709
DB 605 CAGATTACAGAGACACACCTTCTGATGCTTCTTCTAGGGGACCCACACAGACACACACAC 664
QY 1710 GCTGGGCGCCAGGGGAGGAGGACCTTTCAGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1769
DB 665 GCTGGGCGCCAGGGGAGGAGGACCTTTCAGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 724
QY 1770 ATGAGGCTGACACCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1829
DB 725 ATGAGGCTGACACCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 783
QY 1830 CAGAGGTTGTGCGACGCGCTGCGTGAAGTACACCTTCTGAGATCTCTCAGAGACACACTT 1889
DB 784 ATAGGCTGTGTTCACGCGCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 840
QY 1890 GAGATCTGAGGAGACGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1949
DB 841 GAGATCTGAGGAGACGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 899
QY 1950 TAGTCAAACTTTTCTTATCCCGCCCAATCTAT 1985
DB 900 TAGTC-ACACTATTCTTATTCACCACTCTATT 914

RESULT 3
B3722953
LOCUS      B3722953          910 bp      mRNA      linear      F01-06-MAY-2001
DEFINITION      602592618F: NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4824892 5',
                cDNA sequence.
ACCESSION      B3720953
VERSION      B3720953.1  GI:14000140
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  Bases 1 to 910;
  NIH-MGC http://imgc.nhl.nih.gov/;
  National Institutes of Health, Mammalian Gene Collection (MGC);
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  E-mail: rcs@pma.nhl.nih.gov
  Tissue Procurement: Miklos Pakovits, M.D., Ph.D.
  cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
  Toshiyuki and Piero Carninci (RIKEN)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/BLAST at:
  http://image.llnwd.net
  Plate: L24M1C737 row: c column: 05
  High quality sequence steps: 81..
  Location/Qualifiers
    1..910
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /c_son="IMAGE:4824892"
FEATURES
source

```

```

/lab_host="CHI08"
/clone_lib="N:R_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site1: BamHI; Site2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT      213 a  243 c  265 g  189 t
ORIGIN
Query Match      14.1%; Score 459.4; DB 10; Length 910;
Best Local Similarity 76.8%; Pred. No. 4.4e-35;
Matches 650; Conservative 0; Mismatches 186; Indels 10; Gaps 7;
QY 23 GGTGACTAACTACAGAAACGCTGTGGAGGCTCGGGGGGCCAGTSCCGCGT 98
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 GGGGGATTTAACCCAGGAGCGCGCTGTGGAGGCGCGCTGGCGCGCGCATG 63
QY 29 AGTCTCTTCGACCGCGAGCCACACCGCCCGCTGACCATGATAGTGTTCAGGTT 148
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 GCGCTGTCTCGCGCGAGCGCGCACCTACCGAGTGCATGATAGTGTTCAGGTT 123
QY 149 CAACTCCAGCTATGCTTCCAGTGGAGTGTGATCTGACACAGCATCTTCAGGTCGA 208
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
124 CAACTCCAGCATGTTTCCAGTGGAGTGTGATCTGACACAGCATCTTCAGGTCGA 183
QY 209 GGAAGTGGTGTGAAGCAGACAGGGGTTCAGAGCTGACACAGTCTGCTGATTTTGGCGG 268
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 GGGGTGTGTGTAAGCGACAGGGGTTCGCGTGCACAGTTCGCTGTGATTTCCAGG 243
QY 269 GAAGAGCTTCCGATCAGCTGACGGTTCMAAGCTGTGACCTGAGCAACAGAGTATGTT 328
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 GAAGAGCTGAGGATGACTGGACTGTGAGAAATGTGAGCTGGATCAGCAGACGATTGT 303
QY 329 ACACATAGTACAGACACGAGGAGAA--GTGATGAACAAATGATCTGAGGGGAC 386
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
304 TCACATTTGTCAGAGNCGTGGAGCAAAACGTCAGAAATGATGAACCTGGAGCGGAC 353
QY 387 GAACCCAGACACCTCAGAGGGCTTCAATGAGAGTCCAGGACCTTGAACAGTGGAC 446
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 GACCCAGAAAACCGCGGGAGGCTGTGAGCGGAGGCCAGAGCTTGACTCGGGTGGAC 423
QY 447 CTGAGCAGCATACCTCGCGGTGAATCTGTGGGGTGGCGGTGATCTGGACACAGAC 506
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
424 CTCAGCAGCTCAGTCTCCAGGAGCTGTGTGGGGTGGCTGTCACTCTGCACACTGAC 483
QY 507 AGTACAGAGATTCAGACAGCAGAGGTC---CAGTTAAACCACTTACACAGCTTT 563
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
484 AGCAGAGAGATTCACACAGCTGGAAGTCCAGCAGGTAGATCAATCTTACACAGCTTT 543
QY 564 TTATCTACTGCAAGAGGCCCTGCCACAGGTTCAGGCTTGAAAGCTCCGAGTTCAGTGT 623
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
544 TATGTGTATTGTACAGGCCCTGTCAAGAGTGCAGCGGGGAACTCAGGCTACAGTGC 603
QY 624 GGCACCTGCNAACAGACACCTCAGCTTGGCCAGGCCCATCTTCTGGGACGATGTC 683
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
604 AGCAGTGCAGGACGACAGCTCAGCTTGCAGCCCA--GATCCATCTTCTGGGATGATGT 562
QY 684 TTAATTCAAA--CCGATGAGTGGTGAAGTGCAGGCTCCAGAGTCCCTCGGACACGAGC 742
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 TCAATTCNAACCCGATGAGTGGTGAATGTCATCCCACTGTCTCTGGACTAGTGC 722
QY 743 TGAATTTTCTTTA--TATGTGAGACA--CGAAGCTCAGACAGGACAGTGGTAGCT 800
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 AGAATTTTCTTTTACAAATGTGAGCAGACACCCCACTTCTGACGGAACATCAGTAGCT 782
QY 801 TTGAATCTGATCACCAGCAACGAGGCGCAGCATCCCTTGTGATGAGGTGCACA--GATGTGAG 859
D5 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 TTGCACTGTATCCCAACAATACGTCCGGAACATCTCTTCACTTACGTGCACAGGACGTGAG 842

```

QY 860 GAGCCC 865

D5 |||||

843 GAGCCC 848

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY596928 453 bp mRNA linear EST 15-DEC-2002
 BY596928 RIKEN full-length enriched adult inner ear Mus musculus
 CDSa clone F9102K13.3, mRNA sequence.

BY596928 G126932113

EST

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Mus;

1. (bases 1 to 453)

Okazaki, Y., Furuno, K., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Oatso, K., Saito, R., Suzuki, H., Yamakawa, T., Kiyosawa, H.,

Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schorbach, C.,

Sojibuchi, T., Baldarelli, R., Hill, D. P., Eut, C., Hume, D. A.,

Quackenbush, J., Schmitt, L. W., Kanapin, A., Matsuda, H., Batalov, S.,

Beisel, K. W., Blake, C. A., Bradt, D., Brusio, V., Chothia, C., Corbani,

A. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,

A. A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,

Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,

Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzietaki, R. M.,

King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,

P. A., Maglott, D. R., Valtais, L., Marchionni, L., McKenzie, J., Miki,

H., Nagashima, T., Kumata, K., Okido, T., Pavani, K. C., Pertea, G.,

Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. J., Qi, D.,

Ramachandran, S., Pavasi, T., Reed, J. C., Reed, J. J., Reid, J., Ring,

B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempie, C. A., Soccu,

M. B., Shimada, K., Suttara, P., Takenaka, Y., Taylor, M. S., Teasdale,

A. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, S.,

Watanabe, Y., Wells, C., Wilming, L. C., Wyszynski, A., Yanagisawa,

M., Yang, Z., Yuan, Z., Zavolan, M., Zhu, Y., Zierler, A.,

Carninci, P., Hayashizaki, N., Hirozane-Kishikawa, T., Kohno, H., Nakamura,

M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, C., Aizawa, K.,

Y., Itoh, K., Kikawa, T., Kiyazaki, A., Hashizume, M., Tachikawa, K.,

K., Shinnagawa, A., Yasunishi, A., Yoshino, M., Watanabe, K., Watanabe,

E. S., Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-571 (2002)

22154683

22466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome.res@gsr.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Aizawa, T., Carninci, P., Fukuda, S., Hirozane,

K., Imoto, K., Ishii, Y., Itoh, K., Kawai, J., Kohno, H., Miyazaki, A.,

Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, K., Sakai, K.,

Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,

X., Waki, K., Watanabe, K., Muramatsu, M. and Hayashizaki, Y. Direct

Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multiplexed sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Kirk W. Baise: Boys Town National Research Hospital 555 North 30th Street Omaha, NE 68131 USA whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	Location/Qualifiers
source	1..453
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="F930:02K13"
	/tissue_type="inner ear"
	/dev_stage="adult"
	/clone_lib="RIKEN full-length enriched, adult inner ear"
BASE COUNT	122 a 99 c 98 g 134 t
ORIGIN	
Query Match.	13.5% Score 438.8; DB 13; Length 453;
Best Local Similarity	99.3%; Pred. No. 5.8e-33;
Matches 451; Conservative	0; Mismatches 2; Indels 1; Gaps 1;
QY 2775	TCAGGAGCAGTCAGGAGTCGTATGTCACCCATCACTATTTTTCATCATGTCGCT 2834
DB	
QY 2835	GCCTCTCCCTCTCCACACCCGCTGTGAGCTATCGCTATCGGAGCGACGACAAATGTT 2894
DB	
QY 60	GCCTCTGCCTCTCCACACCCGCTGTGAGCTATCGGAGCGACGACAAATGTT 119
DB	
QY 2895	GCTGTTCTGCTTTGGGTAAAGGACCTCCCTGAAGCTGTGGCTCTCCAGTATGTCCT 2954
DB	
QY 120	GCTGTCTGCTTTGGGTAAAGGACCTCCCTGAAGCTGTGGCTCTCCAGTATGTCCT 179
DB	
QY 2955	TTTCCTCTCCACAGATGATATGTTTCTTCAGAAATACATAGTAGTATTTAAATAC 3014
DB	
QY 180	TTTCCTCTCCACAGATGATATGTTTCTTCAGAAATACATAGTAGTATTTAAATAC 239
DB	
QY 3015	CCAAAGACAGGCATCCACAGTGTGTGAGCATGAATCCAGCTGCATTTGTGTGAGTGTG 3074
DB	
QY 240	CCAAAGACAGGCATCCACAGTGTGTGAGCATGAATCCAGCTGCATTTGTGTGAGTGTG 299
DB	
QY 3075	ATAGTGGGATAAGTGTGATGTCAGAGAGTGGAAATCAACCTCTGCAAGCAATCTT 3134
DB	
QY 300	ATAGTGGGATAAGTGTGATGTCAGAGAGTGGAAATCAACCTCTGCAAGCAATCTT 319
DB	
QY 3135	TCTCTTCTGTGAGTGTATTAAGAAATACCTGAAATCTGTGTGTGTGGTACCCAGA 3194
DB	
QY 360	TCTCTTCTGTGAGTGTATTAAGAAATACCTGAAATCTGTGTGTGTGGTACCCAGA 419
DB	
QY 3195	CTGTCAATCAATAAGACCCGCTGCAATGAA 3228
DB	
QY 420	CTGTCAATCAATAAGACCCGCTGCAATGAA 453
DB	

RESULT 5
BY592998
LOCUS
DEFINITION
BY592998 RIKEN full-length enriched, adult inner ear Mus musculus
CDNA clone F93002819 3', mRNA sequence.
ACCESSION
BY592998
VERSION
BY592998.1
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

Ekaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus; 1 (bases : to 484)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Boro, H., Kondo, S., Niki, K., Otsu, N., Saito, R., Suzuki, H., Yamazaki, I., Kiyosawa, H., Yagi, K., Tonari, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gotohori, T., Baldarelli, R., Hill, C. P., Butt, C., Hume, D. A., Gachekush, J., Schrim, L. K., Kanapin, A., Maseda, R., Batilov, S., Beisel, K. N., Blake, C. A., Bradt, C., Brusic, V., Crother, C. F., Corbett, J. E., Cousins, S., Dalia, E., Dragan, T. A., Fletcher, C. F., Godzik, A., Gough, J. K., Gasterland, T., Gariboldi, M., Gissi, C., Fodor, C., Gough, J. K., Grimmond, S., Gustincich, S., Hirokawa, M., Jackson, J. C., Jarvis, E. D., Kama, A., Kawai, H., Kawasawa, Y., Kestler, R. M., King, B. L., Konagaya, A., Kurochkin, V., Lea, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Valtas, L., Vatchinn, J., McKernan, L., Miki, H., Nagashima, T., Numata, K., Okado, T., Pavan, M. J., Petosa, G., Rescde, S., Petrovsky, N., Piliavin, R., Pontius, C. D., Qian, J., Rabadan, R. S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sander, A., Schneider, C., Sempere, C. A., Serou, M., Shirada, K., Sutar, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yacagawa, M., Yang, J., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Koshikawa, T., Kondo, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, Y., Aizawa, K., Aizawa, I., Fukuda, S., Hara, A., Heshizume, W., Imoto, K., Ishii, Y., Itoh, M., Kadawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, X., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcripts based on functional annotation of 60,773 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

1246885

CONTACT: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0245, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome.res@gsr.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imoto, K., Ishii, Y., Itoh, M., Kawai, Y., Kondo, H., Miyazaki, A., Nakamura, M., Nakazawa, K., Nomura, K., Nishizaki, K., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, K., Watanabe, S. M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Gene Sequences. Nucleic Acids Res. 29, 671-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated sequence analysis (RISA) system: a 384-format sequencing pipeline with 384 multi-capillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Kirk W. Baise: Boys Town National Research Hospital 555 North 30th Street Omaha, NE 68131 USA whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

Location/Qualifiers
1..484
/organism="Mus musculus"


```

|||||
360 GAAACAATGCTATGAGGGGAGAGACCCAGACAGGTACAGGGCTCATATGGGAS 418
|||||
423 TCAGGAGGTTGACAGAGTGACCTG 449
|||||
419 TCCAGGAGTTGACAGAGTGACCTG 445
|||||

RESULT 7
BY232905
LOCUS BY232905 333 bp mRNA linear EST 10-DEC-2002
DEFINITION BY232905 RIKEN full-length enriched, adult inner ear Mus musculus
cDNA clone F930028B14 5', mRNA sequence.
ACCESSION BY232905
VERSION BY232905.1 GI:26414015
KEYWORDS Mus musculus (house mouse)
SOURCE EST
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 393)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bonci,H., Kondo,S.,
Nikado,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,K.,
Yagi,K., Tomaru,Y., Hasegawa,Y., Nogari,A., Schonbach,C.,
Gojobori,T., Baldarelli,R., Hill,D.P., Bilt,C., Hume,D.A.,
Oackenhush,J., Schriber,M., Kanapin,A., Matsuda,H., Batalov,S.,
Bailey,K.W., Blake,J.A., Bradt,D., Bruscia,V., Chochia,C., Corbani
,A., Fraser,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Forrest
,J., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
,H., Nagashima,T., Numata,K., Okido,T., Pavan,K.J., Perteira,G.,
Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.C., Reid,J., Ring
,B.Z., Ringwald,M., Sanderlin,A., Schreider,C., Sempile,C.A., Setou
,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y.,
Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yaginisawa
,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
Carrinci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
,X., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirczare
,A., Imotani,K., Ishii,Y., Itch,M., Kawai,J., Konno,H., Miyazaki,A.,
Miyata,M., Nakamura,K., Nomura,K., Numazaki,R., Ohno,M., Sakai,K.,
Sakazume,N., Sakaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami
,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission:
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001).
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format

```

```

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1957-1971 (2000).
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Mark W. Peisner, Boys Town National
Research Hospital, 555 North 30th Street, Omaha, NE 68131, USA, whose
assistance we gratefully acknowledge. Please visit our web site
(http://genome.gsc.riken.go.jp) for further details.
Location/Qualifiers
1..393
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7B2/6J"
/db_xref="taxon:10090"
/clone="F930028B14"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"
BASE COUNT 96 a 99 c 123 g 72 t
ORIGIN
Query Match 11.9%; Score 388.8; DB 13; Length 393;
Best Local Similarity 99.2%; Pred. No. 3.6e-28;
Matches 393; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 7 GAGGGGAGGGGAGGAGGCGCTGGATGACTAAACCTGCAGAGAAACGCTGGTGGAGGCTC 66
DB 1 GAGGGGAGGGGAGGAGGCGCTGGATGACTAAACCTGCAGAGAAACGCTGGTGGAGGCTC 60
QY 67 GGGGGGGCGCCAGTGCCTGGCTAGGTCTCTTCGAGCCGCCAGCCAGCCGCGCGTGA 126
DB 61 GGGGGGGCGCCAGTGCCTGGCTAGGTCTCTTCGAGCCGCCAGCCAGCCGCGCGTGA 120
QY 127 CCATGATAGTGTTCGAGGTCCTCACTCCAGCTATGGCTTCCAGTGGAGGTCTGATTCG 186
DB 121 CCATGATAGTGTTCGAGGTCCTCACTCCAGCTATGGCTTCCAGTGGAGGTCTGATTCG 190
QY 187 ACACCAAGCTTTCGAGCTCAAGGAGTGTTCCTAAGCGAGCGGGGTCCAGCTGACC 246
DB 181 ACACCAAGCTTTCGAGCTCAAGGAGTGTTCCTAAGCGAGCGGGGTTCAGCTGACC 240
QY 247 AGTGGGTGTGTTCCTGGGGAGAGAGCTTCGGAATCTGCTGAGGCTTCAAAACTGTG 306
DB 241 AGTGGGTGTGTTCCTGGGGAGAGAGCTTCGGAATCTGCTGAGGCTTCAAAACTGTG 300
QY 307 ACTTGGAAACAGAGATTTATGACATATGATACAGACCCAGCGAGAGAGAGTATGAAA 366
DB 301 ACTTGGAAACAGAGATTTATGACATATGATACAGACCCAGCGAGAGAGAGTATGAAA 360
QY 367 CAATGATCTTGGAGGGAGAGACCCAGACCA 399
DB 361 CAATGATCTTGGAGGGAGAGACCCAGACCA 393

RESULT 8
BY237113
LOCUS BY237113 418 bp mRNA linear EST 10-DEC-2002
DEFINITION BY237113 RIKEN full-length enriched, adult inner ear Mus musculus
cDNA clone F93010X13 5', mRNA sequence.
ACCESSION BY237113
VERSION BY237113.1 GI:26418305
KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 418)

```


M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Kawai, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Amotani, K., Ishii, Y., Itoh, K., Kagawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 12466851
 Contact: Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0845, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, K., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Yamamatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1772 (2000)

Computer-based methods for the mouse full-length cDNA Encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Kirk W. Beisel (Boys Town National Research Hospital, 555 North 30th Street Omaha, NE 68131 USA) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES
 Source
 1. 351
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="F930042F15"
 /tissue_type="inner ear"
 /dev_stage="adult"
 /clone_lib="RIKEN full-length enriched, adult inner ear"

BASE COUNT
 ORIGIN
 83 a 90 c 111 g 67 t
 Query Match 10.7% Score 349.4; DB 13; Length 351;
 Best Local Similarity 99.7%; Pred. No. 2.2e-24;
 Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GAGGGGAAGGGGAGGAGCGCTGGATGACTAAACCTGACAGAAACGCTGGTGGAGGCTC 66
 Db 1 GAGGGGAAGGGGAGGAGCGCTGGATGACTAAACCTGACAGAAACGCTGGTGGAGGCTC 60
 QY 67 GGGCGGGGCGCAGTSCCGCGTAGTCTCTTCGACCCGCGAGCCACACCCCGCCGGTGA 126
 Db 61 GGGCGGGGCGCAGTSCCGCGTAGTCTCTTCGACCCGCGAGCCACACCCCGCCGGTGA 120
 QY 127 CCATGATAGTGTTCGTAGGTTCAACTCCAGCTATGCGTTCCCACTGGAGGTCGATCTG 186
 Db 121 CCATGATAGTGTTCGTAGGTTCAACTCCAGCTATGCGTTCCCACTGGAGGTCGATCTG 180

QY 187 ACACGAGCATCTTCGAGGTCACAGGAGTGGTGTGTTAAGCGACAGCGGGTTCAGTGAAC 246
 Db 181 ACACGAGCATCTTCGAGGTCACAGGAGTGGTGTGTTAAGCGACAGCGGGTTCAGTGAAC 240
 QY 247 AGTGGGTGTGATTTTTCGCGGGAAGGAGCTTCCGAAATCAGCTGACGGTTCAGAACTGTG 306
 Db 241 AGTGGGTGTGATTTTTCGCGGGAAGGAGCTTCCGAAATCAGCTGACGGTTCAGAACTGTG 300
 QY 307 AACTGGAACACAGAGTATTTACACATAGTACAGACACAGGAGGAGGAGNA 357
 Db 301 AACTGGAACACAGAGTATTTACACATAGTACAGACACAGGAGGAGGAGNA 351

RESULT 10
 CB811663

LOCUS
 DEFINITION
 CB811663 454 bp mRNA linear EST 16-MAY-2003
 srpb2-00216-d4 5' mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Rattus norvegicus (Norway rat);
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Angen, Inc
 One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00216 row: d column: 4.

FEATURES
 Source
 1. 404
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="srpb2-00216-d4"
 /tissue_type="prostate tissue"
 /clone_lib="srpb2 (0220)"
 /note="Vector: pSP0PT1; Site 1: SalI; Site 2: NotI; rat prostate normalized double selected poly(A-) mRNA size fraction > 1 kb"

BASE COUNT
 ORIGIN
 101 a 124 c 92 g 82 t 5 others
 Query Match 10.2% Score 333.4; DB 14; Length 404;
 Best Local Similarity 91.8%; Pred. No. 6.4e-23;
 Matches 391; Conservative 0; Mismatches 31; Indels 2; Gaps 2;

QY 507 AGTAAGAGGATTCAGACGACGACGAGGTCAGTTAAACCCACCTACACAGCTTTTC 566
 Db 1 AGCAAGATGACTCAGACGACGACGAGGTCAGTTAAACCCACCTACACAGCTTTTC 60

QY 567 ATCTACTGAAAGGCGCTTCGACAGGTCGAGCTGGAAAGCTCCGAGTTCAGTGTGGC 626
 Db 61 GTCTACTGAAAGGCGCTTCGACAGGTCGAGCTGGAAAGCTCCGAGTTCAGTGTGGC 120
 QY 627 AACTCAAAACGAGCAACCTTCACCTTGGCCGAGGCCCATCTTCTGGGAGCATGTCTTTA 666
 Db 12: AACTCAAAACGAGCAACCTTCACCTTGGCCGAGGCCCATCTTCTGGGAGCATGTCTTTA 180

QY 687 ATTCGAACCGGATGAGTGGTGAGTCCAGTTCAGACTTCCCTGGAACACGAGACTGAA 746
 Db 181 ATTCGAACCGGATGAGTGGTGAGTCCAGTTCAGACTTCCCTGGAACACGAGACTGAA 240
 QY 747 TTTTCTTTAAATGTGGAGCACACCCACCTCAGACAGGACACGCTCGTASCCTTGAC 806
 Db 241 TTTTCTTTAAATGTGGAGCACACCCACCTCAGACAGGACACGCTCGTASCCTTGAC 300


```

526 CACCACAGAGCTCCAGTTAAACCCACCTACACAGC 560
    ||| ||| ||| ||| ||| ||| ||| ||| |||
445 CCAGAGAGACCGACCGGTAGAGGAGCTACACAGC 519

RESULT 13
A1228346/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A1228346
EST725041 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBRCT55 3' end, mRNA sequence.
A1228346
A1228346.1 G:38:2233
EST...
Rattus sp.
Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 430)
Lee,N.H., Glodok,A., Chandra,I., Mason,T.X., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RSET) Catalog & Rat
Gene Index
Unpublished
Contact: Lee, NH
The Institute for Genomic Research
9710, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0209
Email: nhlee@igr.org
Seq primer: M13-21
Location/Qualifiers
1..430
/organism="Rattus sp."
/ncbi_type="mRNA"
/db_xref="taxon:10118"
/cldone="RBSC755"
/cldone_lib="Normalized rat brain, Bento Soares"
/notes="Organi: brain; Vector: pTV3Fac; Site_1: EcoR1;
Site_2: NotI"
112 a -59 c 77 g 122 t

FEATURES
source
Query Match 6.7%; Score 282.4; DP 9; Length 400;
Best local Similarity 65.8%; Pred. No. 4, Use 18;
Matches 339; Conservative 0; Mismatches 51; Indels 5; Gaps 2;

QY 2863 CGTAATCGGATCGGGAAGCCAGAGAAAGTGTCTGTCTGTTGGGTAAAGGAGCTCC 2322
DB 397 CGATATGGATGGGGAAGCCAGAGAAAGTGTCTGTCTGTTGGGTAAAGGAGCTCC 339
QY 2923 CTGAAGCGCTGTGGGTCTCCAGTATGGTGGCCCTTCTCTCTCTAAAGATGCGATGTCTT 3962
DB 337 CTGAAGCGCTCAATGGGCTCTGGTATGGTGGCCCTTCTCTCTCTAAAGATGCGATGTCTT 278
QY 2983 CTTCAGATACATATAGTATATCTCTTAATATAGCGAAGACAGAGGCTCCACAGTGTGTGA 3042
DB 207 CTTCAGATACAGACAGTATCTTAAATATAGTCAAAATGAAGACACCCATAGTGTGCGA 218
QY 3043 GCATGAATACAGCGCTGGCATTTGTGTAGTGTGATAGTGGGATAAAGTGGATGTGAGA 3101
DB 217 GCATGAATACAGCGCTGGCATTTGTGTAGTGTGATAGTGGGATAAAGTGGATGTGAGA 358
QY 3102 AGAGTGGAAATCAAAACCTCTGCAAGCAATCTTTCTTTCTTCTGTGAAGTGTATTAAGAAA 3161
DB 157 AGAGTGGAAATCAAAACCTCTGAAAAG---CTCTTCTCTCTGTGAAGTGTATTAAGAAA 102
QY 3162 TACCTGAGCTCTCTGTGTGTGTGTACCCAGACTGTCAATCAATAAAGACCCAGACTGT 3221
DB 101 TACCTGAGCTCTCTGTGTGTGTGTACCCAGACTGTCAATCAATAAAGTGTGTGTGTGTGT 42
QY 3222 CAATGAAGAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 3255

```

```

Db 41 CAAATCAATAGTACCCAGACTGTCAATAAAGAA 8

RESULT 14
BF286222/c 391 bp mRNA linear EST 28-NOV-2003
LOCUS BF286222 Rat Gene Index, normalized rat, Rattus norvegicus cDNA
DEFINITION Rattus norvegicus cDNA clone RG1FN37 3' sequence, mRNA sequence.
ACCESSION BF286222
VERSION BF286222.1 GI:11217292
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
1 (bases 1 to 391)
AUTHORS Yalcin, R.; Cho, J.; Lee, Y.; Karamycheva, S.; Parvizli, B.; Perle, G.;
Suldana, R.; Tsai, C.; White, J.; Ouderkubus, J. and Lee, N.H.
TITLE Generation of ESTs from Normalized Rat Embryo, Bento Soares
JOURNAL Unpublished
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20852, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information.
FEATURES
Location/Qualifiers
1..391
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/cGene="RG1FN37"
/tissue_type="mixed tissue"
/lab_host="DH5-alpha"
/clone_lib="Rat Gene Index, normalized rat, Rattus
norvegicus cDNA"
/notes="Vector: pTZ19Pac; Site 1: EcoRI; Site 2: NotI;
Combination of ROV, RBR, RKT, RLT, RPL, RLU, REV, RYU, RSP
108 a 89 c 74 g 121 t

BASE COUNT
OrigIn
Query Match 8.4%; Score 273.8; DB 10; Length 391;
Best Local Similarity 85.3%; Pred. No. 3e-17;
Matches 330; Conservative 0; Mismatches 52; Indels 5; Gaps 2;

QY 2870 GCATGGGAAGCAGGACCAATGTTTCTGCTTCTGCTTGGTAAAGGAGCTCCCTGAAGC 2929
DB 390 GCATGGGAAGCAGGACCAATGTTTCTGCTTCTGCTTGGTAAAGGAGCTCCCTGAAGC 331
QY 2930 TCTGTGGCTCTCAGTATGCTCCCTTCTCTTCTTAACAGATGCATATGTTTCTTCTGCA 2989
DB 330 TCATGGCTCTCTGGTATGCTCCCTTCTCTTCTTCTGAGATTGCATATGTTTCTTCTGCA 271
QY 2990 ATACATAGTATGTTCTTAATAATACCCAAAGACAGGATCCACAGTGTGTGAGCATGAA 3049
DB 270 ATACACAGTATGTTCTTAATAATACCTCAATGAAGACACCATCATGTGTGGAGCATGAA 211
QY 3050 TCACAGCTGCTGTTGTGAGTGTGTAATAGTGGAT-AAAATGATGTGTGAGAGAGTGG 3108
DB 210 TCACAGCTGCTGTTGTGAGTGTGTAATAGTGGAT-AAAATGATGTGTGAGAGAGTGG 151
QY 3109 AAATCAACCTCTGCAAGCAATCTTCTTCTGTTGAGTGTATTAAGATATCTGCA 3166
DB 150 AAATCAACCTTGTGAAAG-CTTCTTCTCTCTGTTGAGTGTATTAAGATATCTGCA 95
QY 3169 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3225
DB 94 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 35

```

```

QY 3229 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3255
DB 34 TAAGTACCCAGACTGTCAATAAAGAA 8

RESULT 15
BF649338 274 bp mRNA linear EST 04-SEP-2003
LOCUS BF649338
DEFINITION UI-X-BH2.3-acc-f-08-0-UI.r1 NIH BMAP M.S3.3 Mus musculus cDNA clone
ACCESSION UI-X-BH2.3-acc-f-08-0-UI.5, mRNA sequence.
VERSION BF649338.1 GI:9975162
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 274)
AUTHORS Borralco, X.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (5), 791-806 (1996)
MEDLINE 87044477
PUBMED 8999548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 2706
Fax: 301 443 9895
Email: EST@mail.nih.gov
CNA Library Preparation: M.B. Soares Lab Clone Distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seg primer: V3 Reverse
Location/Qualifiers
1..274
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cGene="UI-X-BH2.3-acc-f-08-0-UI"
/GenStage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH-BMAP_M.S3.3"
/notes="Vector: pTZ19Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco R1; The
NIH-BMAP_M.S3.3 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs, pineal
hypothalamus, cortex, amygdala, basal ganglia, pineal
ganglion, striatum, hippocampus) after a series of
subtractions to reduce the representation of cDNAs from
which ESTs had already been generated. The following
serially subtracted libraries were generated in this
process: NIH-BMAP_M.S3.3, NIH-BMAP_M.S2, NIH-BMAP_M.S1.
The subtracted library (NIH-BMAP_M.S3.3) was constructed
as follows: PCR-amplified cDNA inserts from NIH-BMAP_M.S2
clones from which 3' ESTs had been derived was used as a
driver in a hybridization with the NIH-BMAP_M.S2 library
in the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the
NIH-BMAP_M.S3.3 library. This procedure has been
previously described (Borralco, Lennon and Soares, Genome
Research 6:791-806, 1996)."
55 a 76 c 82 g 59 t
BASE COUNT

```

ORIG: N

```

Query Match      8.4%; Score 272.4; DB 10; Length 274;
Best Local Similarity 99.6%; Pred. No. 5.5e-17;
Matches 273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 38 AACCTGACAGAAACGCTGGTGGAGGCTCGGGCGGGCCAGTGCCTCCGTAGGTCCTTC 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 AACCTGACAGAAACGCTGGTGGAGGCTCGGGCGGGCCAGTGCCTCCGTATGTCCTTC 60

QY 98 TCGACCCGAGCCACACCCCGCCCGGTGACCATGATGATGTTGTGAGGTTCAACTCCAG 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TCGACCCGAGCCACACCCCGCCCGGTGACCATGATGATGTTGTGAGGTTCAACTCCAG 120

QY 158 CTATGGCTTCCAGTGGAGGTGGATTTGACACCAAGCATCTTGCAGCTCAAGGAAGTGGT 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CTATGGCTTCCAGTGGAGGTGGATTTGACACCAAGCATCTTGCAGCTCAAGGAAGTGGT 180

QY 218 TGGTAAGCGACAGGGGTTCCAGCTGACCAAGCTGCGTGTGATTTTTCGGGGAAGGAGCT 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TGC'AAGCGACAGGGGTTCCAGCTGACCAAGCTGCGTGTGATTTTTCGGGGAAGGAGCT 240

QY 278 TCCGAATCACCTGACGTTCAAACTGTGACCTG 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TCCGAATCACCTGACGTTCAAACTGTGACCTG 274

```

Search completed: October 24, 2003, 13:32:05
 Job time : 4330 secs

2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
1988
1987
1986
1985
1984
1983
1982
1981
1980
1979
1978
1977
1976
1975
1974
1973
1972
1971
1970
1969
1968
1967
1966
1965
1964
1963
1962
1961
1960
1959
1958
1957
1956
1955
1954
1953
1952
1951
1950
1949
1948
1947
1946
1945
1944
1943
1942
1941
1940
1939
1938
1937
1936
1935
1934
1933
1932
1931
1930
1929
1928
1927
1926
1925
1924
1923
1922
1921
1920
1919
1918
1917
1916
1915
1914
1913
1912
1911
1910
1909
1908
1907
1906
1905
1904
1903
1902
1901
1900

```
XX PT New polynucleotides encoding mouse parkin2 protein, useful for
PT producing a transgenic non-human animal as an animal model for
XX neurodegenerative diseases
XX FS Disclosure; Page 17-19; 62pp; English.
XX CC The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans.
CC The human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the
CC efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes including
CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
CC stroke, vascular irregularities or metabolic irregularities, associated
CC with a less active or non-active parkin protein.
XX SQ Sequence 464 AA;
Query Match 100.0%; Score 2570; DB 22; Length 464;
Best Local Similarity 100.0%; Pred. No. 2,6e-236;
Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MIVFVRNSSYGFFVEVDSDTSLQLKEVAKRGVPADQLRVIFAGKELPNHLLTVQNC 60
DB 1 MIVFVRNSSYGFFVEVDSDTSLQLKEVAKRGVPADQLRVIFAGKELPNHLLTVQNC 60
QY 61 LEQGSIVHIVORRRRSHETNASGDPQSTSEGSIMESRLTRVDSLSHTLPVDSVGLA 120
DB 61 LEQGSIVHIVORRRRSHETNASGDPQSTSEGSIMESRLTRVDSLSHTLPVDSVGLA 120
QY 121 VILDTGKRDSEARGPVKPTYNSEFFYCKGPKHVQKLVQCGCKQATLTLAGGPPS 180
DB 121 VILDTGKRDSEARGPVKPTYNSEFFYCKGPKHVQKLVQCGCKQATLTLAGGPPS 180
QY 181 CWDDVLIPINRMSGCGQPCDPTRAEFKCGAHTSDKDTISVALNLTTSNRRSIPCIAC 240
DB 181 CWDDVLIPINRMSGCGQPCDPTRAEFKCGAHTSDKDTISVALNLTTSNRRSIPCIAC 240
QY 241 TDVRSPLVFOCNHRHVICDCEFLHYCVTRLNDQFVHDAQLGYSLPCVAGCPNLSIKEL 300
DB 241 TDVRSPLVFOCNHRHVICDCEFLHYCVTRLNDQFVHDAQLGYSLPCVAGCPNLSIKEL 300
QY 301 HHFRILGEEQVTRYQYQYGAEECVLQMGVLCPRGCGAGLLPEQGQRKVTCEGNGLGCG 360
DB 301 HHFRILGEEQVTRYQYQYGAEECVLQMGVLCPRGCGAGLLPEQGQRKVTCEGNGLGCG 360
QY 361 FVFCRDCKEAYHEGDCDLSLEPSPGATSCQAYRVDKRAAEQARWEASKEITKTKPCPRC 420
DB 361 FVFCRDCKEAYHEGDCDLSLEPSPGATSCQAYRVDKRAAEQARWEASKEITKTKPCPRC 420
QY 421 NVPIEKNGGCHMKCPQPOCKLEWNCWNGCGENRACNGCHWFDV 464
DB 421 NVPIEKNGGCHMKCPQPOCKLEWNCWNGCGENRACNGCHWFDV 464
RESULT 2
AAB67531
ID AAB67531 standard; Protein: 464 AA.
AC AAB67531;
XX 29-YAV-2001 (first entry)
DT XX Amino acid sequence of a mutated murine parkin2 polypeptide.
DE XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
```

```
KW metabolic irregularity.
XX Mls sp.
XX EP1081225-A1.
XX 07-MAR-2001.
XX 30-AUG-1999; 99EP-0116766.
XX 30-AUG-1999; 99EP-0116766.
XX (BIOF-) BIOFRONTIERA PHARM. GMBH.
XX Liebbert H;
XX MPI; 2001-212797722.
XX N-PSDB; AAF55258.
XX New polynucleotides encoding mouse parkin2 protein, useful for
XX producing a transgenic non-human animal as an animal model for
XX neurodegenerative diseases
XX Claim 7; Page 47-49; 62pp; English.
XX The present sequence represents a murine parkin2 polypeptide. The
XX sequence contains the mutation Lys61Asn. Mutations or deletions in
XX the parkin2 gene cause Parkinson's disease in humans. The human
XX parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
XX and polynucleotides are useful for analysing neurodegenerative diseases.
XX They are also useful for testing the efficacy of the treatment of a
XX neurodegenerative disease such as Parkinson's disease, Alzheimer's
XX disease, Huntington's disease, amyotrophic lateral sclerosis,
XX Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease,
XX and secondary causes including Parkinson's syndromes like toxins, drugs,
XX brain tumours, head trauma, stroke, vascular irregularities or metabolic
XX irregularities, associated with a less active or non-active parkin
XX protein.
XX SQ Sequence 464 AA;
Query Match 99.6%; Score 2565; DB 22; Length 464;
Best Local Similarity 99.6%; Pred. No. 7.9e-236;
Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MIVFVRNSSYGFFVEVDSDTSLQLKEVAKRGVPADQLRVIFAGKELPNHLLTVQNC 60
DB 1 MIVFVRNSSYGFFVEVDSDTSLQLKEVAKRGVPADQLRVIFAGKELPNHLLTVQNC 60
QY 61 LEQGSIVHIVORRRRSHETNASGDPQSTSEGSIMESRLTRVDSLSHTLPVDSVGLA 120
DB 61 LEQGSIVHIVORRRRSHETNASGDPQSTSEGSIMESRLTRVDSLSHTLPVDSVGLA 120
QY 121 VILDTGKRDSEARGPVKPTYNSEFFYCKGPKHVQKLVQCGCKQATLTLAGGPPS 180
DB 121 VILDTGKRDSEARGPVKPTYNSEFFYCKGPKHVQKLVQCGCKQATLTLAGGPPS 180
QY 181 CWDDVLIPINRMSGCGQPCDPTRAEFKCGAHTSDKDTISVALNLTTSNRRSIPCIAC 240
DB 181 CWDDVLIPINRMSGCGQPCDPTRAEFKCGAHTSDKDTISVALNLTTSNRRSIPCIAC 240
QY 241 TDVRSPLVFOCNHRHVICDCEFLHYCVTRLNDQFVHDAQLGYSLPCVAGCPNLSIKEL 300
DB 241 TDVRSPLVFOCNHRHVICDCEFLHYCVTRLNDQFVHDAQLGYSLPCVAGCPNLSIKEL 300
QY 301 HHFRILGEEQVTRYQYQYGAEECVLQMGVLCPRGCGAGLLPEQGQRKVTCEGNGLGCG 360
DB 301 HHFRILGEEQVTRYQYQYGAEECVLQMGVLCPRGCGAGLLPEQGQRKVTCEGNGLGCG 360
QY 361 FVFCRDCKEAYHEGDCDLSLEPSPGATSCQAYRVDKRAAEQARWEASKEITKTKPCPRC 420
DB 361 FVFCRDCKEAYHEGDCDLSLEPSPGATSCQAYRVDKRAAEQARWEASKEITKTKPCPRC 420
```

QY 421 NVPIEKNGGCHMKCPQCKLEWNCGCCWNRACMGDHFV 464
 DB 421 NVPIEKNGGCHMKCPQCKLEWNCGCCWNRACMGDHFV 464

RESULT 3
 AAB67532
 ID AAB67532 standard; Protein; 464 AA.
 AC AAB67532;
 XX
 XX 29-MAY-2001 (first entry)
 XX
 XX Amino acid sequence of a mutated murine parkin2 polypeptide.
 XX
 XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW Brain tumor; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.
 XX
 CS Mus sp.
 XX
 XX EP1091225-A1.
 XX
 XX 07-MAR-2001.
 XX
 XX 30-AUG-1999; 99EP-0116766.
 XX
 XX 30-AUG-1999; 99EP-0116766.
 XX
 XX (BIOF-) BIOFRONTIERA PHARM GMBH.
 XX
 XX Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR N-PSDB; AAF55259.
 XX
 XX New polynucleotides encoding mouse parkin2 protein, useful for
 PT producing a transgenic non-human animal as an animal model for
 PT neurodegenerative diseases
 XX
 XX Claim 7; Page 49-51; 62pp; English.
 XX
 XX The present sequence represents a murine parkin2 polypeptide. The
 CC sequence contains the mutation Thr415Asn. Mutations or deletions in
 CC the parkin2 gene cause Parkinson's disease in humans. The human
 CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
 CC and polynucleotides are useful for analysing neurodegenerative diseases.
 CC they are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis,
 CC Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease,
 CC and secondary causes inducing Parkinson's syndromes like toxins, drugs,
 CC brain tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.
 XX
 XX Sequence 464 AA;
 SQ

Query Match 99.8%; Score 2565; DB 22; Length 464;
 Best Local Similarity 99.8%; Pred No. 7, 8e-236;
 Matches 463; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIVVAFNFSYGFPEVDSDTILQLKEWAKRQGVADQLRVIPAGKELPNHVTQKCD 60
 DB 1 MIVVAFNFSYGFPEVDSDTILQLKEWAKRQGVADQLRVIPAGKELPNHVTQKCD 60
 QY 61 LQOQS*VHVIVQVRPRRSHTNAGSGDEQSTSEGSWESRS*TRVDLSSTHTPVDSVGLA 120
 DB 61 LQOQS*VHVIVQVRPRRSHTNAGSGDEQSTSEGSWESRS*TRVDLSSTHTPVDSVGLA 120
 QY 121 VILDTSKSDSEAAAGPVKPTYNSEFFYCKGPCHKVQPKLRVCGTKQATLLAQGPS 180

DB 121 VILDTSKSDSEAAAGPVKPTYNSEFFYCKGPCHKVQPKLRVCGTKQATLLAQGPS 180
 QY 181 QKCDVLPNRMSEGCQSPDCPGTRAEEFFKCGAHTSDKDTISVALNLTISNRSIPICAC 240
 DB 181 QKCDVLPNRMSEGCQSPDCPGTRAEEFFKCGAHTSDKDTISVALNLTISNRSIPICAC 240
 QY 241 TDVRSPLVFCQNRHNVICLDCFFHYCVTRINRPFVHDAQLGYSDVCVAGCQNSLIKEL 300
 DB 241 TDVRSPLVFCQNRHNVICLDCFFHYCVTRINRPFVHDAQLGYSDVCVAGCQNSLIKEL 300
 QY 301 KHFRILGSEQYTRYQXGABECVLMQGVLCPTPGCGAGLTPCQSGKVTCEGNGSLGCG 360
 DB 301 KHFRILGSEQYTRYQXGABECVLMQGVLCPTPGCGAGLTPCQSGKVTCEGNGSLGCG 360
 QY 361 FVFCRDCKEAVHEGDCSLLPEPSGATSCAYEVKRAAEQARWEASKETIKKTKPCPRC 420
 DB 361 FVFCRDCKEAVHEGDCSLLPEPSGATSCAYEVKRAAEQARWEASKETIKKTKPCPRC 420
 QY 421 NVPIEKNGGCHMKCPQCKLEWNCGCCWNRACMGDHFV 464
 DB 421 NVPIEKNGGCHMKCPQCKLEWNCGCCWNRACMGDHFV 464

RESULT 4
 AAB67533
 ID AAB67533 standard; Protein; 451 AA.
 XX
 XX AAB67533;
 XX
 XX 29-MAY 2001 (first entry)
 XX
 XX Amino acid sequence of a mutated murine parkin2 polypeptide.
 XX
 XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW Brain tumor; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.
 XX
 CS Mus sp.
 XX
 XX EP1091225-A1.
 XX
 XX 07-MAR-2001.
 XX
 XX 30-AUG-1999; 99EP-0116766.
 XX
 XX 30-AUG-1999; 99EP-0116766.
 XX
 XX (BIOF-) BIOFRONTIERA PHARM GMBH.
 XX
 XX Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR N-PSDB; AAF55260.
 XX
 XX New polynucleotides encoding mouse parkin2 protein, useful for
 PT producing a transgenic non-human animal as an animal model for
 PT neurodegenerative diseases
 XX
 XX Claim 7; Page 51-53; 62pp; English.
 XX
 XX The present sequence represents a murine parkin2 polypeptide. The
 CC sequence contains the mutation Trp433Stop. Mutations or deletions in
 CC the parkin2 gene cause Parkinson's disease in humans. The human
 CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
 CC and polynucleotides are useful for analysing neurodegenerative diseases.
 CC they are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis,
 CC Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease,
 CC and secondary causes inducing Parkinson's syndromes like toxins, drugs,
 CC brain tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.
 XX
 XX Sequence 451 AA;
 SQ

```

CC brain tumours, head trauma, stroke, vascular irregularities or metabolic
CC irregularities, associated with a less active or non-active parkin
CC protein.
XX
SQ Sequence 451 AA;
Query Match: 96.6%; Score 2483; DB 22; Length 451;
Best Local Similarity 100.0%; Pred. No. 5e-228;
Matches 451; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
QY 1 MIVVFRNSSYGFPEVDSQTSILQKEVAKQGVPAQQLRVIFAGKELPNHLITVQNC 60
DQ 1 MIVVFRNSSYGFPEVDSQTSILQKEVAKQGVPAQQLRVIFAGKELPNHLITVQNC 60
QY 61 LEQOSIVHIVQRPFRSHETNASGGDEPOSTSGS:WESRSLTRVDLSHTLIPVDSVGLA 120
DQ 61 LEQOSIVHIVQRPFRSHETNASGGDEPOSTSGS:WESRSLTRVDLSHTLIPVDSVGLA 120
QY 121 VILDTSKRDEAARGPVKPTYNSTFIYCKGPKRVQCGTKQATLTLAQGPS 180
DQ 121 VILDTSKRDEAARGPVKPTYNSTFIYCKGPKRVQCGTKQATLTLAQGPS 180
QY 181 CWDCVILPNRMSEGCSPDCPTRAEFFPKCGAHTSDKDTSVALLNITSNRSPICAC 240
DQ 181 CWDCVILPNRMSEGCSPDCPTRAEFFPKCGAHTSDKDTSVALLNITSNRSPICAC 240
QY 241 TDVRSVLPVFCNHRHVICLDCPHLYCVTRLNDRQFVHDAQGLGYSLPVAGPNSLIKEL 300
DQ 241 TDVRSVLPVFCNHRHVICLDCPHLYCVTRLNDRQFVHDAQGLGYSLPVAGPNSLIKEL 300
QY 301 HHFRILGEGEYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGGKVTCEGNGLGCG 360
DQ 301 HHFRILGEGEYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGGKVTCEGNGLGCG 360
QY 361 FVFCRCCKAYHEGDCDLSLEPSGATSOAYRVYDKRAAEAGARWEAEASKETIKTTKPCPRC 420
DQ 361 FVFCRCCKAYHEGDCDLSLEPSGATSOAYRVYDKRAAEAGARWEAEASKETIKTTKPCPRC 420
QY 421 NVPIEKNGGCMHKKCPQCKLEWCMNCGCE 451
DQ 421 NVPIEKNGGCMHKKCPQCKLEWCMNCGCE 451
RESULT 5
AAE30801
ID AAE30801 standard; Protein: 465 AA.
XX
AC AAE30801;
XX
XX 24-FEB-2003 (first entry)
DE Human Parkin D126 mutant protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischemic stroke; Parkinson's disease; Alzheimer's disease; neurotrophic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KW mutant; muten.
XX
OS Homo sapiens.
CS Synthetic.
XX
FH Key Location/Qualifiers
FT Cleavage-site 126
FT Misc-difference 126
FT /note= "Wild-type Asp is replaced with Glu"
FT Misc-difference 223
FT /note= "Encoded by CCA"
XX
PN NC020279459-A2.
XX
XX 10-OCT-2002.
XX
XX 02-APR-2002; 2002MO-DK00221.

```

```

XX 29-MAR-2001; 2001DK-0000525.
XX 03-APR-2001; 2001US-282288F.
XX (NSGE-1) NSGENE AS.
XX Jensen PH;
XX
XX KFI: 2003-046812/04.
XX N-PSDB; AAD47630.
XX
XX New isolated nucleic acid sequence encoding a Parkin polypeptide,
XX useful for treating, preventing or diagnosing neurological disorders,
XX e.g. Parkinson's disease, Alzheimer's disease or ischemic stroke, and
XX in screening assays.
XX
XX Claim 10; Page 71; Tipp: English.
XX
XX The invention relates to Parkin protein and its corresponding nucleic
XX acid sequence. The nucleic acid sequence is useful for altering the
XX proteolytic processing of Parkin at its potential cleavage site at Asp
XX 126. The invention is used in manufacturing or testing a pharmaceutical.
XX composition for treating and/or preventing a neurological disorder, e.g.
XX Alzheimer's disease or ischemic stroke. It also used for detecting the
XX occurrence of proteolytic processing of Parkin at Asp 126 in a sample,
XX in monitoring a potential disposition for a neurodegenerative disease,
XX and for treating, preventing and/or diagnosing Parkinson's disease or
XX other neurodegenerative disorders. The viral vector is used for
XX transforming neuronal cells in vivo or ex vivo. The invention is useful
XX for neurotransplantation into the CNS of a mammal. It may be used in
XX screening assays to identify compounds that increase or decrease
XX apoptosis. It is also used in gene therapy. The present sequence is
XX human Parkin mutant protein.
XX
SQ Sequence 455 AA;
Query Match: 94.5%; Score 2171.5; DB 24; Length 465;
Best Local Similarity 83.4%; Pred. No. 2.8e-193;
Matches 388; Conservative 30; Mismatches 46; Indels 1; Gaps 1;
QY 1 MIVVFRNSSYGFPEVDSQTSILQKEVAKQGVPAQQLRVIFAGKELPNHLITVQNC 60
DQ 1 MIVVFRNSSYGFPEVDSQTSILQKEVAKQGVPAQQLRVIFAGKELPNHLITVQNC 60
QY 61 LEQOSIVHIVQRPFRSHETNASGGDEPOSTSGS:WESRSLTRVDLSHTLIPVDSVGLA 120
DQ 61 LEQOSIVHIVQRPFRSHETNASGGDEPOSTSGS:WESRSLTRVDLSHTLIPVDSVGLA 120
QY 121 VILDTSKRDEAARGPVKPTYNSTFIYCKGPKRVQCGTKQATLTLAQGP 179
DQ 121 VILDTSKRDEAARGPVKPTYNSTFIYCKGPKRVQCGTKQATLTLAQGP 180
QY 180 SCWDVILPNRMSEGCSPDCPTRAEFFPKCGAHTSDKDTSVALLNITSNRSPICAC 239
DQ 180 SCWDVILPNRMSEGCSPDCPTRAEFFPKCGAHTSDKDTSVALLNITSNRSPICAC 240
QY 240 CTDVRSVLPVFCNHRHVICLDCPHLYCVTRLNDRQFVHDAQGLGYSLPVAGPNSLIKE 299
DQ 240 CTDVRSVLPVFCNHRHVICLDCPHLYCVTRLNDRQFVHDAQGLGYSLPVAGPNSLIKE 300
QY 300 LHFRILGEGEYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGGKVTCEGNGLGCG 359
DQ 300 LHFRILGEGEYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGGKVTCEGNGLGCG 363
QY 360 GFVFCRCCKAYHEGDCDLSLEPSGATSOAYRVYDKRAAEAGARWEAEASKETIKTTKPCPR 413
DQ 360 GFVFCRCCKAYHEGDCDLSLEPSGATSOAYRVYDKRAAEAGARWEAEASKETIKTTKPCPR 420
QY 420 CNVPIEKNGGCMHKKCPQCKLEWCMNCGCEANRACMGDHFV 464
DQ 421 CHVPEKNGGCMHKKCPQCKLEWCMNCGCEANRACMGDHFV 465

```

```

RESULT 6
AAV32501
ID AAV32501 standard; Protein: 465 AA.
XX AC
XX AAV32501;
XX DT 21-OCT-1999 (first entry)
XX DE Human parkin gene variant protein.
XX KW Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX CS Homo sapiens.
XX ZN W09940191-A1.
XX PD 12-AUG-1999.
XX PF 09-FEB-1999; 99WC-JPC0545.
XX PR 09-FEB-1999; 98JP-0027531.
XX PA (MIZU) MIZUNO Y.
XX FA (SHIMIZU) SHIMIZU N.
XX PI Mizuno Y, Shimizu N;
XX WP 1999-494295/41.
XX DR N-PSDB; AAX99923.
XX PT Gene implicated in the pathology of Parkinson's disease, used for treatment of the disease
XX PS Claim 1; Page 83-88; 114pp; English.
XX CC This sequence is encoded by a gene of the invention, and is implicated in the pathology of Parkinson's disease. This sequence is a variant of the parkin gene found in parkinson's disease patients. The sequences may be used for the diagnosis, treatment (including gene therapy) and investigation of Parkinson's disease.
XX SQ Sequence 465 AA;
Query March 84.5%; Score 2170.5; DB 20; Length 465;
Best Local Similarity 83.4%; Pred. No. 3.5e-198;
Matches 388; Conservative 29; Mismatches 47; Indels 1; Gaps 1;
QY 1 MIVFVRFNSYGFVEYDSDTSILQLKEVYVAKRQGVADQLRVIYFAGKELPNHJTVQKCD 60
DQ 1 MIVFVRFNSYGFVEYDSDTSILQLKEVYVAKRQGVADQLRVIYFAGKELPNHJTVQKCD 60
QY 61 LEOQSIYVHVORPRRSHEHETNAGSGDEPSTSEGSINERSLTRVDLSSTLPVDSVGLA 120
DQ 61 LEOQSIYVHVORPRRSGEENATGGDDPRNAAGGCEBPQSILTRVDLSSTLVPGDSVGLA 120
QY 121 VILDTSDKRDSEARGPV-KPTVNSPFIYCKGPCHKVQPGKLRVCGCTCKATLTLAGGP 179
DQ 121 VILDTSDKRDSPAGSPAGRSIYNSVYVCKPGQVQPGKLRVCGCTCKATLTLAGGP 180
QY 180 SCDDVLIPNMGEGQSPDCTRAFFPKCAHPTSDKDTSYALNLTITNRSIPGIA 239
DQ 181 SCDDVLIPNMGEGQSPDCTRAFFPKCAHPTSDKDTSYALNLTITNRSITCIT 240
QY 240 CTDVRSFVJVFQGNHRIHVICDGFHLYCYVTRNDROFVHDAOLGYSLPCVAGCPNSLIKE 299
DQ 241 CTDVRSFVJVFQGNHRIHVICDGFHLYCYVTRNDROFVHDAOLGYSLPCVAGCPNSLIKE 300
QY 300 LHFPRILGEOYTRYQYGAEECVLQXGVLCPRPCGAGLLPEQGRKVTCEGNGLGC 359
DQ 301 LHFPRILGEOYTRYQYGAEECVLQXGVLCPRPCGAGLLPEQGRKVTCEGNGLGC 360
QY 360 GFVFCROCKEAYHEGDCDSLLPEFGATSCAYRVCKEAAECAEKAEASKETIKTKPCPR 419

```

```

DQ 361 GFVFCROCKEAYHEGDCDSAVFEASGTTTQAYRVDERAAEQARNEAASKETIKTKPCPR 420
QY 420 CNVPIEKNGGCMHKCPQPCCKLEWNCWCENRACVGEHWFEDV 464
DQ 421 CHVPVEKNGGCMHKCPQPCCKLEWNCWCENRACVGEHWFEDV 465
RESULT 7
AAE30800
ID AAE30800 standard; Protein: 465 AA.
XX AC
XX AAE30800;
XX DT 24-FEB-2003 (first entry)
XX DE Human Parkin protein.
XX KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy; ischaemic stroke; Parkinson's disease; Alzheimer's disease; neurotrophic; transgenic; cerebroprotective; neuroprotective; neurotransplantation.
XX CS Homo sapiens.
XX PH Key Location/Qualifiers
XX F1 Cleavage-site 126
XX F2 WC200279459-A2.
XX F3 10-OCT-2002.
XX F4 02-APR-2002; 2002WO-DK00221.
XX F5 29-MAR-2001; 2001DK-0003525.
XX F6 03-MAR-2001; 2001US-281284P.
XX F7 INSGE-1 NSGENE AS.
XX F8 Jensen PR;
XX F9 WPI; 2003-046812/04.
XX F10 N-PSDB; AAD47679.
XX PT New isolated nucleic acid sequence encoding a parkin polypeptide, useful for treating, preventing or diagnosing neurological disorders, e.g. Parkinson's disease, Alzheimer's disease or ischemic stroke, and in screening assays.
XX PS Claim 10; Page 69; 71pp; English.
XX CC The invention relates to parkin protein and its corresponding nucleic acid sequence. The nucleic acid sequence is useful for altering the proteolytic processing of parkin at its potential cleavage site at Asp 126. The invention is used in manufacturing or testing a pharmaceutical composition for treating and/or preventing a neurological disorder, e.g. Alzheimer's disease or ischaemic stroke. It also used for detecting the occurrence of proteolytic processing of parkin at Asp 126 in a sample, in monitoring a potential disposition for a neurodegenerative disease, and for treating, preventing and/or diagnosing parkinson's disease or other neurodegenerative disorders. The viral vector is used for transforming neuronal cells in vivo or ex vivo. The invention is useful for neurotransplantation into the CNS of a mammal. It may be used in screening assays to identify compounds that increase or decrease apoptosis. It is also used in gene therapy. The present sequence is human parkin protein.
XX SQ Sequence 465 AA;
Query March 84.5%; Score 2170.5; DB 24; Length 465;
Best Local Similarity 83.4%; Pred. No. 3.5e-198;
Matches 388; Conservative 29; Mismatches 47; Indels 1; Gaps 1;
QY 1 MIVFVRFNSYGFVEYDSDTSILQLKEVYVAKRQGVADQLRVIYFAGKELPNHJTVQKCD 60

```



```

Db      1  MIVFVRNNSHGFPVEVDSTSLQLEKVVAKQGVPAQDLRVIFAGKELRNDWTVCNC 60
QY      61  LEQOSIVHIYVCRPRRSHETHNASGGDEPQSTSGSWEKSLTRVLDLSHTLPVDSVGLA 120
Db      62  LDOOSIVHIYVCRPRRSHETHNASGGDEPQSTSGSWEKSLTRVLDLSHTLPVDSVGLA 120
QY      121  VILDTDSKRSDEAARGPV-KPTVNSFFYCKGPKCHKVPQKLVQCGTCKQATLILACGP 179
Db      122  VILDTDSKRSDEAARGPV-KPTVNSFFYCKGPKCHKVPQKLVQCGTCKQATLILACGP 180
QY      180  SCMDVLLPNRMGECQSPDCPGTRAEFFPKCGAHPSTSDKDTSVLNLITSNRRSPICIA 239
Db      181  SCMDVLLPNRMGECQSPDCPGTRAEFFPKCGAHPSTSDKDTSVLNLITSNRRSPICIA 240
QY      240  CTDVRSVLVFOCNRHNVICLDCFHLVYCVTRLNDRQFVHDQALGYSLPCVAGCPNSLIKE 299
Db      241  CTDVRSVLVFOCNRHNVICLDCFHLVYCVTRLNDRQFVHDQALGYSLPCVAGCPNSLIKE 300
QY      300  LHHFRILGEEQYTRYQOYGAECEVLMGGVLCPRPGCGAGLLPEQGRKVTCEGNGLGC 359
Db      301  LHHFRILGEEQYTRYQOYGAECEVLMGGVLCPRPGCGAGLLPEQGRKVTCEGNGLGC 360
QY      360  GFVFCRCKEAYHEGDCDSJLEBPSGATSOAYRYVDKRAAEQARWEAEASKETIKKTKPCPR 419
Db      361  GFAPCRCKEAYHEGDCDSJLEBPSGATSOAYRYVDKRAAEQARWEAEASKETIKKTKPCPR 420
QY      420  CNVPIEXNGGGMHMKCPQPCCKLEWNCNCCENRACMGSHWFDV 464
Db      421  CNVPIEXNGGGMHMKCPQPCCKLEWNCNCCENRACMGSHWFDV 465

```

RESULT 8

AAV32502

ID AAV32502 standard; Protein: 437 AA.

XX XX

AC AAV32502;

DT 21-OCT-1999 (first entry)

DE Human parkin gene variant protein.

XX XX Parkinson's disease related gene; parkin gene; variant; gene therapy.

XX OS Homo sapiens.

XX PN W09940191-A1.

XX PC 12-AUG-1999.

XX PF 09-FEB-1999; 99WO-JP00545.

XX PR 09-FEB-1998; 98JP-0027531.

XX PA (MIZU); MIZUNO Y.

XX PA (SHIM); SHIMIZU N.

XX PI Mizuno Y, Shimizu N;

XX DR WPI; 1999-494295/41.

XX DR N-PSDB; AAX99924.

XX PT Gene implicated in the pathology of Parkinson's disease, used for treatment of the disease

XX XX Claim 1; Page 89-94; 114pp; English.

XX CC This sequence is encoded by a gene of the invention, and is implicated in the pathology of Parkinson's disease. This sequence is a variant of the parkin gene found in Parkinson's disease patients. The sequences may be used for the diagnosis, treatment (including gene therapy) and investigation of Parkinson's disease.

XX XX Sequence 437 AA;

Query Match

Best Local Similarity 77.6%; Score 1993.5; DB 20; Length 437;

Matches 362; Conservative 29; Mismatches 45; Indels 23; Gaps 2;

QY 1 MIVFVRNNSHGFPVEVDSTSLQLEKVVAKQGVPAQDLRVIFAGKELRNDWTVCNC 60

DB 1 MIVFVRNNSHGFPVEVDSTSLQLEKVVAKQGVPAQDLRVIFAGKELRNDWTVCNC 60

QY 61 LEQOSIVHIYVCRPRRSHETHNASGGDEPQSTSGSWEKSLTRVLDLSHTLPVDSVGLA 120

DB 61 LEQOSIVHIYVCRPRRSHETHNASGGDEPQSTSGSWEKSLTRVLDLSHTLPVDSVGLA 120

QY 121 VILDTDSKRSDEAARGPV-KPTVNSFFYCKGPKCHKVPQKLVQCGTCKQATLILACGP 179

DB 121 VILDTDSKRSDEAARGPV-KPTVNSFFYCKGPKCHKVPQKLVQCGTCKQATLILACGP 179

QY 180 SCMDVLLPNRMGECQSPDCPGTRAEFFPKCGAHPSTSDKDTSVLNLITSNRRSPICIA 239

DB 179 SCMDVLLPNRMGECQSPDCPGTRAEFFPKCGAHPSTSDKDTSVLNLITSNRRSPICIA 239

QY 240 CTDVRSVLVFOCNRHNVICLDCFHLVYCVTRLNDRQFVHDQALGYSLPCVAGCPNSLIKE 299

DB 240 CTDVRSVLVFOCNRHNVICLDCFHLVYCVTRLNDRQFVHDQALGYSLPCVAGCPNSLIKE 299

QY 300 LHHFRILGEEQYTRYQOYGAECEVLMGGVLCPRPGCGAGLLPEQGRKVTCEGNGLGC 359

DB 273 LHHFRILGEEQYTRYQOYGAECEVLMGGVLCPRPGCGAGLLPEQGRKVTCEGNGLGC 332

QY 360 GFVFCRCKEAYHEGDCDSJLEBPSGATSOAYRYVDKRAAEQARWEAEASKETIKKTKPCPR 419

DB 333 GFAPCRCKEAYHEGDCDSJLEBPSGATSOAYRYVDKRAAEQARWEAEASKETIKKTKPCPR 392

QY 420 CNVPIEXNGGGMHMKCPQPCCKLEWNCNCCENRACMGSHWFDV 464

DB 393 CNVPIEXNGGGMHMKCPQPCCKLEWNCNCCENRACMGSHWFDV 437

RESULT 9

AAB67521

ID AAB67521 standard; Protein: 344 AA.

XX XX

AC AAB67521;

CT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine truncated parkin2 polypeptide.

XX XX

KW Parkinson's disease; 6q25.2-27; neurodegenerative disease;

KW Alzheimer's disease; Huntington's disease; atrophic lateral sclerosis;

KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

KW Brain tumour; Head trauma; stroke; vascular irregularity;

XX XX metabolic irregularity.

XX OS Mus sp.

XX XX

XX PN EPI08-225-A1.

XX XX

XX PC 07-MAR-2001.

XX XX

XX PF 30-AUG-1999; 99EP-0116766.

XX XX

XX PR 30-AUG-1999; 99EP-0116766.

XX XX

XX PA (BIOF); BIOFRONTIERA PHARM GMBH.

XX XX

XX PI Luebbert H;

XX XX

XX DR WPI; 2001-212797/22.

XX DR N-PSDB; AAF55249.

XX XX

XX PT New polynucleotides encoding mouse parkin2 protein, useful for producing a transgenic non-human animal as an animal model for

PT neurodegenerative diseases -
 XX Claim 7; Page 40-41; 62pp; English.
 XX The present sequence represents a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains a deletion, leading to a truncated
 CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
 CC disease in humans. The human parkin2 gene is located in gene region
 CC 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
 CC analysing neurodegenerative diseases. They are also useful for testing
 CC the efficacy of the treatment of a neurodegenerative disease such as
 CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 CC Pick's disease, Prion disease, and secondary causes including
 CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
 CC stroke, vascular irregularities or metabolic irregularities, associated
 CC with a less active or non-active parkin protein.
 XX Sequence 344 AA;
 SQ

Query Match 72.6%; Score 1857; DB 22; Length 344;
 Best Local Similarity 74.1%; Pred. No. 2,1e-165;
 Matches 344; Conservative C; Mismatches C; Indels 120; Gaps 1;

QY 1 MIVFVFNSSYGPFVEVDSSTSLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
 DB 1 MIVFVFNSSYGPFVEVDSSTSLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 56
 QY 61 LEQOSIVHIVQRRRRSHETNAGSGDEPQSTSEGSWESESLRVDLSHSLPVDVSGLA 120
 DB 57 -----QGPS 56
 QY 121 VILDTDSKDSSEARGPVKPTNSFFIYCKGPKHVQPKLRVQCGTKQATLTACQPS 180
 DB 57 -----QGPS 60
 QY 181 CWDDVLIPNRMSEGCSPDCPTRAEFFKCGAPTSKDTSVALLNLTSSRRSPICAC 240
 DB 181 CWDDVLIPNRMSEGCSPDCPTRAEFFKCGAPTSKDTSVALLNLTSSRRSPICAC 120
 QY 241 TVRSPVLVFCQNRHNVICLDCFHLYCVTRLNDRQFVHCAQLGYSLPCV 289
 DB 241 TVRSPVLVFCQNRHNVICLDCFHLYCVTRLNDRQFVHCAQLGYSLPCV 180
 QY 301 HFRILGEGEYTRYQYGAECEVLQGGVLCPRPGCGAGLLPEQGRKYTCGGKGLCG 240
 DB 301 HFRILGEGEYTRYQYGAECEVLQGGVLCPRPGCGAGLLPEQGRKYTCGGKGLCG 360
 QY 361 FVFCRCCKEAYHEGDCDSLLPESGATSOAYRVYDKRAEQARWEASKETIKKTAKPCERC 420
 DB 241 FVFCRCCKEAYHEGDCDSLLPESGATSOAYRVYDKRAEQARWEASKETIKKTAKPCERC 300
 QY 421 NVPIENGGGMHMKQPOCKLEWCNCGCENWRACMGDHFV 464
 DB 301 NVPIENGGGMHMKQPOCKLEWCNCGCENWRACMGDHFV 344

RESULT 10
 AAB67526
 ID AAB67526 standard; Protein: 296 AA.
 XX AAB67526;
 XX 29-MAY-2001 (first entry);
 XX Amino acid sequence of a murine truncated parkin2 polypeptide.
 XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 XX brain tumour; head trauma; stroke; vascular irregularity;
 XX metabolic irregularity.

OS Mus sp.
 PK EP1081225-A1.
 PC 07-MAR-2001.
 PF 30-AUG-1999; 99SP-0116766.
 PR 30-AUG-1999; 99EP-0116766.
 PA (BIOF-1) BIOFRONTERA PHARM GMBH.
 XX Liebbert H;
 XX XP1: 2001-212797/22.
 DR N-PS23; AAF55253.
 XX New polynucleotides encoding mouse parkin2 protein, useful for
 PT producing a transgenic non-human animal as an animal model for
 FC neurodegenerative diseases -
 XX Claim 7; Page 44-45; 62pp; English.
 XX The present sequence represents a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains a deletion, leading to a truncated
 CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
 CC disease in humans. The human parkin2 gene is located in gene region
 CC 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
 CC analysing neurodegenerative diseases. They are also useful for testing
 CC the efficacy of the treatment of a neurodegenerative disease such as
 CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 CC Pick's disease, Prion disease, and secondary causes including
 CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
 CC stroke, vascular irregularities or metabolic irregularities, associated
 CC with a less active or non-active parkin protein.
 XX Sequence 296 AA;
 SQ

Query Match 60.3%; Score 1550; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 3,2e-139;
 Matches 289; Conservative C; Mismatches C; Indels C; Gaps C;

QY 1 MIVFVFNSSYGPFVEVDSSTSLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
 DB 1 MIVFVFNSSYGPFVEVDSSTSLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
 QY 61 LEQOSIVHIVQRRRRSHETNAGSGDEPQSTSEGSWESESLRVDLSHSLPVDVSGLA 120
 DB 61 LEQOSIVHIVQRRRRSHETNAGSGDEPQSTSEGSWESESLRVDLSHSLPVDVSGLA 120
 QY 121 VILDTDSKDSSEARGPVKPTNSFFIYCKGPKHVQPKLRVQCGTKQATLTACQPS 180
 DB 121 VILDTDSKDSSEARGPVKPTNSFFIYCKGPKHVQPKLRVQCGTKQATLTACQPS 180
 QY 181 CWDDVLIPNRMSEGCSPDCPTRAEFFKCGAPTSKDTSVALLNLTSSRRSPICAC 240
 DB 181 CWDDVLIPNRMSEGCSPDCPTRAEFFKCGAPTSKDTSVALLNLTSSRRSPICAC 240
 QY 241 TVRSPVLVFCQNRHNVICLDCFHLYCVTRLNDRQFVHCAQLGYSLPCV 289
 DB 241 TVRSPVLVFCQNRHNVICLDCFHLYCVTRLNDRQFVHCAQLGYSLPCV 289

RESULT 11
 AAB67519
 ID AAB67519 standard; Protein: 250 AA.
 XX AAB67519;
 XX 29-MAY-2001 (first entry);
 XX Amino acid sequence of a murine parkin2 polypeptide.

```

XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
CS Mus sp.
XX
XX EP1081225-A1.
XX
XX 07-MAR-2001.
XX
XX 30-AUG-1999; 99EP-0116766.
XX
XX 30-AUG-1999; 99EP-0116766.
XX
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX
XX Luebbert H;
XX
XX WPI; 2001-212797/22.
XX
XX N-PSDB; AAF55246.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
PT producing a transgenic non-human animal as an animal model for
PT neurodegenerative diseases -
XX
XX Claim 7; Page 21-22; 62pp; English.
XX
XX The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans.
CC The human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the
CC efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes including
CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
CC stroke, vascular irregularities or metabolic irregularities, associated
CC with a less active or non-active parkin protein.
XX
SQ Sequence 250 AA;
Query Match 49.9%; Score 1281.5; DB 22; Length 250;
Best Local Similarity 99.6%; Pred. No. 1.1e-113;
Matches 244; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MIVFVRFNSYSGFPVEVDSSTLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
DB 1 MIVFVRFNSYSGFPVEVDSSTLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
QY 61 LEQQSIVHIVQRPFRSHETNASGDEPOSTSEGSINWESRSLTRVDLSHSLPVDVSGLA 120
DB 61 LEQQSIVHIVQRPFRSHETNASGDEPOSTSEGSINWESRSLTRVDLSHSLPVDVSGLA 120
QY 121 VILDTSKSDSEAAARGP-VKPTVNSFFIYCKGPKHKVQGLRVQCGTKQATLTLAQQP 179
DB 121 VILDTSKSDSEAAARGP-VKPTVNSFFIYCKGPKHKVQGLRVQCGTKQATLTLAQQP 180
QY 180 SCWDDVLIIPNRMSEGCSPDCPGTTRAEFFPKCGAHP-TSDKQTSVALNLTNSRRESIPCIA 239
DB 181 SCWDDVLIIPNRMSEGCSPDCPGTTRAEFFPKCGAHP-TSDKQTSVALNLTNSRRESIPCIA 240
QY 240 CTQVR 244
DB 241 CTQVR 245

```

RESULT 12

AAB67518

ID AAB67518 standard; Protein; 262 AA.

```

XX AAB67518;
XX
XX 23-MAY-2001 (first entry)
XX
XX Amino acid sequence of a murine parkin2 polypeptide.
XX
XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
XX brain tumour; head trauma; stroke; vascular irregularity;
XX
XX metabolic irregularity.
XX
XX Mus sp.
XX
XX EP1081225-A1.
XX
XX 07-MAR-2001.
XX
XX 30-AUG-1999; 99EP-0116766.
XX
XX 30-AUG-1999; 99EP-0116766.
XX
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX
XX Luebbert H;
XX
XX WPI; 2001-212797/22.
XX
XX N-PSDB; AAF55245.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
PT producing a transgenic non-human animal as an animal model for
PT neurodegenerative diseases -
XX
XX Claim 7; Page 19-20; 62pp; English.
XX
XX The present sequence represents a murine parkin2 polypeptide. Mutations
CC or deletions in the parkin2 gene cause Parkinson's disease in humans.
CC The human parkin2 gene is located in gene region 6q25.2-27. Parkin2
CC polypeptides and polynucleotides are useful for analysing
CC neurodegenerative diseases. They are also useful for testing the
CC efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes including
CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
CC stroke, vascular irregularities or metabolic irregularities, associated
CC with a less active or non-active parkin protein.
XX
SQ Sequence 262 AA;
Query Match 49.7%; Score 1276.5; DB 22; Length 262;
Best Local Similarity 99.6%; Pred. No. 1.1e-113;
Matches 243; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MIVFVRFNSYSGFPVEVDSSTLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
DB 1 MIVFVRFNSYSGFPVEVDSSTLQLKEVAKRGQVPADQLRVIFAGKELPNHLTVQNC 60
QY 61 LEQQSIVHIVQRPFRSHETNASGDEPOSTSEGSINWESRSLTRVDLSHSLPVDVSGLA 120
DB 61 LEQQSIVHIVQRPFRSHETNASGDEPOSTSEGSINWESRSLTRVDLSHSLPVDVSGLA 120
QY 121 VILDTSKSDSEAAARGP-VKPTVNSFFIYCKGPKHKVQGLRVQCGTKQATLTLAQQP 179
DB 121 VILDTSKSDSEAAARGP-VKPTVNSFFIYCKGPKHKVQGLRVQCGTKQATLTLAQQP 180
QY 180 SCWDDVLIIPNRMSEGCSPDCPGTTRAEFFPKCGAHP-TSDKQTSVALNLTNSRRESIPCIA 239
DB 181 SCWDDVLIIPNRMSEGCSPDCPGTTRAEFFPKCGAHP-TSDKQTSVALNLTNSRRESIPCIA 240
QY 240 CTQVR 243

```

Db 241 CTDV 244

RESULT 13
ID ABB65114 standard; Protein; 468 AA.
XX
XX ABB65114;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 22134.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmacological.
XX
XX Drosophila melanogaster.
XX
XX W0200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US092331.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 23-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PM, Myers EM;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; AB09217.
XX
XX New isolated nucleic acid detection reagent for detecting 100C or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 22134; 21pp * Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 100C or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL16175), expressed DNA
CC sequences (ABL16184C-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIP
CC at: ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 468 AA;
PS Query Match 41.9%; Score 1078; DB 22; Length 468;
Best Local Similarity 43.7%; Pred. No. 6,7e-94;
Matches 229; Conservative 78; Mismatches 145; Indels 46; Gaps 12;
XX
XX 3 VVRNNSVGFVEVDSTSLQLEKVVAKQGVPAQDLRVIPAGKELPHLTVQVQDLE 62
Db 18 EVKNTGKTLVNLLEPQMDKNVKELVAPQGLQDPLKIIIPAGKELSDATTEQCDLG 77
XX
XX 63 QGSIVHIV-RP-RRRHETNAGSDPEQSTSEGSIVESRSLTRVDSLSTLPVDSVGL 119
XX
XX 78 QGVJHAIHLRPPVCRQKQKQATLEEPSLSDAS-----KPLNETLC 121
XX
XX 120 AVILTDISK---RDSARGPVKPTNSFFTYGKPKVCPKLRVQCTQCATFLIA 176
XX
XX 122 ELQLESEERLNITDEERVEAKA----HFFVHC-SQCDKLCNGKLRVRCALCKGAFVH 175
XX
XX 177 QGPSVQGVLIENRMSGCCSPD--CPGTR-----AEFFFKCAHPT-SDKTSVALNL 227
XX
XX 176 RDEPCWQGVKGRIFGHGCESEVACVENVAGDPFPAEFFFKCAHVSQGGKDFAAFLNL 235

228 ITSNRESICIACTDVRSPVLVFCQNRHVVICLDCFELHYCVTRINDRCFVHDAOLGYSLP 287
236 IKNIKNVCLACTVSVTLVFPCCASQHVVICDCFRHYCRSRIGERQFYPHFPGTILP 295
238 QVAGCFNSJIKELHFRILGEEYTRYQYGAEECVLQNGVLCPRFGCGAGALPEGQR 347
296 CPAGCESPIETTHPKLLTREYDPRYGRFATEEVVLQAGGVLCPCPCGNGVILVEPCR 356
345 KVTCEGNGJGGGFVFCRCCKEAVHESGCDLLHSIGAT---SCAYRVEKRAAFQASWE 404
356 KVTQCN---GGYVFCRKCCLQYHIGEC---DFEGTGASATNSCEYTVDPVFAAEARW 409
405 ASKETIKTKTKPCPRKNVTEYGGCMHYKTPQCKLEWCKWCKCCBNRACMGDHW 462
410 ASNVTKKVTCKPCPKCTPTEDGGCMHVMCTRAGCGFENQWCVQTEKDCNKAHWF 467

RESULT 14
AAB67525
CC AAB67525 standard; Protein; 183 AA.
XX
XX AAB67525;
XX
XX 29-MAY-2001 (first entry)
XX
XX Amino acid sequence of a murine truncated parkin2 polypeptide.
XX
XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW Brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
XX Mus sp.
XX
XX EPI081225-A1.
XX
XX 07-MAR-2001.
XX
XX 30-AUG-1999; 99EP-0116765.
XX
XX 30-AUG-1999; 99EP-0116766.
XX
XX (BIOP-) BOPRONTERA PHARY GMER.
XX
XX Luebbert H;
XX
XX KP; 2001-212737/22.
XX
XX N-PSDB; AAF55252.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for
PT producing a transgenic non-human animal as an animal model for
PT neurodegenerative diseases.
XX
XX Claim 7; Page 43-44; 62pp; English.
XX
XX The present sequence represents a murine parkin2 polypeptide. The
CC polynucleotide sequence contains a deletion, leading to a truncated
CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
CC disease in humans. The human parkin2 gene is located in gene region
CC 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
CC analysing neurodegenerative diseases. They are also useful for testing
CC the efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes including
CC Parkinson's syndromes like toxins, drugs, brain tumours, head trauma,
CC stroke, vascular irregularities or metabolic irregularities, associated
CC with a less active or non-active parkin protein.
XX
XX Sequence 183 AA;
SQ

```

Best Local Similarity 100.0%; Pred. NO. 4,6e-73; Indels 0; Gaps 0;
Matches 177; Conservative 0; Mismatches 0;
QY 1 MIVVAFNFSYGPPVEVDCSTSIQLQKEVYAKQGVPAQQLRVIFAGKELPNHETVQNCQ 60
      |||||
Db 1 MIVVAFNFSYGPPVEVDCSTSIQLQKEVYAKQGVPAQQLRVIFAGKELPNHETVQNCQ 60
      |||||
QY 61 LEQCSIVHIVCRPRRSHETNASGGDEPOSTSESSIWESRSLTRVLSLSHTLPVDSVGLA 120
      |||||
Db 61 LEQCSIVHIVCRPRRSHETNASGGDEPOSTSESSIWESRSLTRVLSLSHTLPVDSVGLA 120
      |||||
QY 121 VILDTQKRDSEAAAGPVKFTYNSFIYKGGPHKVKQPKLRVCCGTCQKATLTAAQ 177
      |||||
Db 121 VILDTQKRDSEAAAGPVKFTYNSFIYKGGPHKVKQPKLRVCCGTCQKATLTAAQ 177
      |||||

```

```

Search completed: October 21, 2003, 21:28:46
Job time : 87 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ON protein - protein search, using sw model

Run on: October 21, 2003, 21:27:25 ; Search time 20 Seconds
(without alignments)
981.622 Million cell updates/sec

Title: US-09-830-703-4

Perfect score: 2570

Sequence: 1 MIVPVRFNSSYGFVEVDS.....CWNCCENRACMGDHWFV 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 326717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/2/iaa/6C COMB.pep:*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	261.5	10.2	503	4	US-09-914-259-13 Sequence 13, Appl
2	186	7.2	587	1	US-08-398-008A-2 Sequence 2, Appl
3	186	7.2	587	2	US-08-393-233-2 Sequence 2, Appl
4	136	5.3	352	3	US-08-834-764-2 Sequence 2, Appl
5	136	5.3	352	5	PCT-US95-09377-2 Sequence 2, Appl
6	134	5.2	1121	1	US-07-789-915A-2 Sequence 2, Appl
7	134	5.2	1121	1	US-08-005-002C-2 Sequence 2, Appl
8	134	5.2	1121	1	US-08-487-203A-2 Sequence 2, Appl
9	129.5	5.0	158	4	US-09-091-725-25 Sequence 25, Appl
10	128	5.0	128	2	US-08-557-309B-41 Sequence 41, Appl
11	128	5.0	128	3	US-08-834-306-41 Sequence 41, Appl
12	128	5.0	128	3	US-08-993-674A-41 Sequence 41, Appl
13	128	5.0	128	4	US-09-256-976-41 Sequence 41, Appl
14	127.5	5.0	101	4	US-09-482-611B-99 Sequence 99, Appl
15	127.5	5.0	101	4	US-09-482-611B-101 Sequence 101, Appl
16	126	4.9	76	1	US-08-232-815-2 Sequence 2, Appl
17	126	4.9	76	1	US-08-350-936-2 Sequence 2, Appl
18	126	4.9	76	5	PCT-US95-04536-2 Sequence 2, Appl
19	125.5	4.9	103	2	US-08-771-201-9 Sequence 35, Appl
20	124	4.9	75	1	US-08-350-884-35 Sequence 35, Appl
21	124	4.8	75	1	US-08-709-773-35 Sequence 35, Appl
22	124	4.8	75	2	US-08-709-177-35 Sequence 35, Appl
23	124	4.8	75	2	US-08-833-678A-3 Sequence 3, Appl
24	124	4.8	75	3	US-08-529-169A-3 Sequence 3, Appl
25	124	4.8	75	3	US-09-483-799-3 Sequence 3, Appl
26	124	4.8	76	1	US-08-450-834-6 Sequence 6, Appl
27	124	4.8	76	4	US-09-516-277-6 Sequence 6, Appl

28	124	4.8	77	6	5510474-3 Patent No. 5510474
29	124	4.8	78	2	US-08-505-486-94 Sequence 94, Appl
30	124	4.8	78	3	US-08-801-028-94 Sequence 94, Appl
31	124	4.8	78	3	US-09-340-154-94 Sequence 94, Appl
32	124	4.8	78	4	US-09-482-611B-94 Sequence 100, Appl
33	124	4.8	78	4	US-09-482-611B-102 Sequence 102, Appl
34	124	4.8	78	4	US-09-482-611B-102 Sequence 94, Appl
35	124	4.8	78	5	PCT-US95-09336-94 Sequence 94, Appl
36	124	4.8	78	5	PCT-US95-09336-94 Sequence 2, Appl
37	124	4.8	457	4	US-09-516-277-2 Sequence 4, Appl
38	124	4.8	457	4	US-09-516-277-4 Sequence 2, Appl
39	124	4.8	533	3	US-08-452-092-2 Sequence 2, Appl
40	124	4.8	533	3	US-08-746-622-2 Sequence 2, Appl
41	124	4.8	533	3	US-09-094-350-2 Sequence 2, Appl
42	124	4.8	533	6	5510474-2 Patent No. 5510474
43	124	4.8	1242	4	US-09-488-270A-2 Sequence 2, Appl
44	123.5	4.8	76	1	US-08-466-274-2 Sequence 2, Appl
45	122.5	4.8	2813	3	US-09-896-449A-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-914-259-13
Sequence 13, Application US/09814259
Patent No. 6495336
GENERAL INFORMATION:
APPLICANT: Yakowski, Lee
APPLICANT: Hyman, Paul
APPLICANT: Williams, Mark
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
FILE REFERENCE: 847-010-999
CURRENT APPLICATION NUMBER: US/09/914.259
CURRENT FILING DATE: 2000-11-21
NUMBER OF SEQ ID NOS: 180
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 503
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-914-259-13

Query Match:	10.2%	Score 261.5;	DB 4;	Length 503;
Best Local Similarity:	28.8%	Pred. No. 1.4e-18;		
Matches	79;	Conservative	37;	Mismatches 97; Indels 61; Gaps 16;
Q1	208	FRKCGAHTSDKDTSVANLITSNRBS:PCIACTDVRSP--VLVPQCNHSHVLCJCCFHL	265	
D1	105	FFKC-AKVNPPNATAIKOKTSRSQCECE:FSQLPDSMAGLECHHR--FCXPCWHS	161	
Q2	266	YCVTRLMDRCFVHDAQIGYSLPCA-GCPNSL:KELHFRILGEEQY-TRYQYGAEECV	323	
D2	162	YLSTKI-----VAEG-LGQTISCAHSC-DILVDDVTVAKLVTDARVYKVQCLITNSFV	214	
Q3	324	LQMGVV--GPRGCG-AGLLPEQGRKVTCSGNGLGGCFVFCRCKEAYHS-----	373	
D3	215	-ECNQLLRWCSVDCYAVKVPVAEPYVCK-----GQVFCFACGNGWHPKCRAL	267	
Q4	374	----GCDSSLRESGATSOARVCKAAECARHERASKETIKKTKPCPCPCNVPIEKNG	429	
D4	269	KYIKKCDP-----DSSTSW-----IAANTKECPCCSVTIKCGS	303	
Q5	430	CMHKCPQPOCKLEWKNWNGCENRACMGDHWF	463	
D5	304	CNHVCKNCKNEFCNVCGLGSEF--HSSSWVN	335	

RESULT 2
US-09-398-008A-2
Sequence 2, Application US/08398008A
Patent No. 5665588
GENERAL INFORMATION:

APPLICANT: Kornbluth, Jacki
 TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gilbreath & Adler, P.C.
 STREET: 8011 Candle Lane
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 COMPUTER: MACINTOSH IICx
 OPERATING SYSTEM: Macintosh
 SOFTWARE: Microsoft Word 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/398,008A
 FILING DATE: March 2, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/126,501
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Adler, Dr. Benjamin Aaron
 REGISTRATION NUMBER: 35,423
 REFERENCE/DOCKET NUMBER: 05705CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (713) 777-2322
 TELEFAX: (713) 777-6908
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 587
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 HYPOTHETICAL: No
 ANTI-SENSE: No
 US-08-398-008A-2

Query Match 7.2%; Score 186; DB 1; Length 587;
 Best Local Similarity 25.5%; Pred. No. 1.4e-10;
 Matches 85; Conservative 28; Mismatches 104; Indels 116; Gaps 23;
 QY 176 AGPSCWDVLPNRMSEGCSPDGPTRABF-PFKCGAHTSDKDTSVALLNLTSHRRS 234
 DB 72 ACGPP-----PEALPAE-----PAAEAAGAAAGAGPGPDDEAEGG-GPGGEE 116
 QY 235 IPCIAC-----TDVRSPVLVFCNHRHVICLDCFHY-----CVTRINDR 274
 DB 117 VECPLCVLRPPERAPRLI-SCPHS--CRCLRHLYLRLSESRVPSICEPESERLNP- 172
 QY 275 QFVHDAGLQVSLPCVAGCPNSJIKELHHF---RILGEEQYTRYQCYGAEVCVLMGGVLC 331
 DB 173 ---HDRLALLAP-----PLMHXYEEFIVRRYLASDPCRW-----C 206
 QY 332 PRGGGAGLL-----PECGORKVTCGGNGLCGFVFCDCKEAYHEGD-CDSILLES 383
 DB 207 PAPDCGYAVIAYGCASCP-----KUTCERE--GCQTEFCHCKQWHPNOTCVMARQ-- 256
 QY 384 GATSOAYRV-DKAAEQARWEASKEITKTKPCPCRCNVPIEK--NGGGMHMKCPQPPQ 440
 DB 257 -QRAQTLRVRTKHTSGLSYGQSGSPDDI-----KPCPRCSAYIIKMNDSCHNHTC--AVC 309
 QY 441 KLEWCMNC-----GC-----ENNR 454
 DB 310 GCEFCWLCMKETSLLHLYSPSGCTFWGKKPWSR 342

RESULT 3
 US-08-893-333-2

Sequence 2, Application US/08893333A
 Patent No. 5981705
 GENERAL INFORMATION:
 APPLICANT: Kornbluth, Jacki
 TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
 FILE REFERENCE: 05705CIP/C
 CURRENT APPLICATION NUMBER: US/08/893,333A
 CURRENT FILING DATE: 1997-07-16
 NUMBER OF SEQ ID NOS: 17
 SEQ ID NO 2
 LENGTH: 587
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Amino acid sequence of Natural Killer Lytic
 OTHER INFORMATION: Associated Protein encoded by nucleotides
 OTHER INFORMATION: 190 to 1953 of Sequence ID. No. 5981705 1
 Patent No. 5981705
 US-08-893-333-2

Query Match 7.2%; Score 186; DB 2; Length 587;
 Best Local Similarity 25.5%; Pred. No. 1.4e-10;
 Matches 85; Conservative 28; Mismatches 104; Indels 116; Gaps 23;
 QY 176 AGPSCWDVLPNRMSEGCSPDGPTRABF-PFKCGAHTSDKDTSVALLNLTSHRRS 234
 DB 72 ACGPP-----PEALPAE-----PAAEAAGAAAGAGPGPDDEAEGG-GPGGEE 116
 QY 235 IPCIAC-----TDVRSPVLVFCNHRHVICLDCFHY-----CVTRINDR 274
 DB 117 VECPLCVLRPPERAPRLI-SCPHS--CRCLRHLYLRLSESRVPSICEPESERLNP- 172
 QY 275 QFVHDAGLQVSLPCVAGCPNSJIKELHHF---RILGEEQYTRYQCYGAEVCVLMGGVLC 331
 DB 173 ---HDRLALLAP-----PLMHXYEEFIVRRYLASDPCRW-----C 206
 QY 332 PRGGGAGLL-----PECGORKVTCGGNGLCGFVFCDCKEAYHEGD-CDSILLES 383
 DB 207 PAPDCGYAVIAYGCASCP-----KUTCERE--GCQTEFCHCKQWHPNOTCVMARQ-- 256
 QY 384 GATSOAYRV-DKAAEQARWEASKEITKTKPCPCRCNVPIEK--NGGGMHMKCPQPPQ 440
 DB 257 -QRAQTLRVRTKHTSGLSYGQSGSPDDI-----KPCPRCSAYIIKMNDSCHNHTC--AVC 309
 QY 441 KLEWCMNC-----GC-----ENNR 454
 DB 310 GCEFCWLCMKETSLLHLYSPSGCTFWGKKPWSR 342

RESULT 4
 US-08-894-764-2
 Sequence 2, Application US/08894764
 Patent No. 6103500
 GENERAL INFORMATION:
 APPLICANT: Innes, Michael
 TITLE OF INVENTION: Production of Tissue Factor Pathway
 TITLE OF INVENTION: Inhibitor
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton St.
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/894,764

```

; FILING DATE: 12-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/286,510
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0991.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-28-854-764-2

Query Match 5.3%; Score 136; DB 3; Length 352;
Best Local Similarity 25.6%; Pred. No. 1.1e-05;
Matches 44; Conservative 28; Mismatches 44; Indels 56; Gaps 6;

QY 1 MIVFVFNSSYGFPVEVDSCTSIQLKEVVAKGVADQRLVIFAGKELPHLTVNCD 60
DB 1 MOIFVKTLTGKTTLEVESSTIDNVKSKIQCKEGFPDQQLIFAGKQLEDGRTLSQYN 60
QY 61 LEQGSIVHIVQPRRSRSHETNAGSGDPOSTSEGSIMESRSILTRVDLSHSLTVDSVGLA 120
DB 61 IQKESLHLVLR-----GGDSEED-----EHT----- 85
QY 121 VLDLTD-----SKDSEAAAGPVKPTVNSF-----FY--CKG 151
DB 86 IITDTEPLPLKMHSCAFKADGPKA:WKAPFFNIPTRQEEIYGGCEG 137

RESULT 5
PCT-US95-09377-2
; Sequence 2, Application PC/TUS9509377
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION
; TITLE OF INVENTION: Production of Tissue Factor Pathway
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent'n Release #1.0, Version #1.10B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09377
; FILING DATE: 25-JULY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 0991.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2585
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
; PCT-US95-09377-2

Query Match 5.3%; Score 136; DB 5; Length 352;
Best Local Similarity 25.6%; Pred. No. 1.1e-05;
Matches 44; Conservative 28; Mismatches 44; Indels 56; Gaps 6;

QY 1 MIVFVFNSSYGFPVEVDSCTSIQLKEVVAKGVADQRLVIFAGKELPHLTVNCD 60
DB 1 MOIFVKTLTGKTTLEVESSTIDNVKSKIQCKEGFPDQQLIFAGKQLEDGRTLSQYN 60
QY 61 LEQGSIVHIVQPRRSRSHETNAGSGDPOSTSEGSIMESRSILTRVDLSHSLTVDSVGLA 120
DB 61 IQKESLHLVLR-----GGDSEED-----EHT----- 85
QY 121 VLDLTD-----SKDSEAAAGPVKPTVNSF-----FY--CKG 151
DB 86 IITDTEPLPLKMHSCAFKADGPKA:WKAPFFNIPTRQEEIYGGCEG 137

RESULT 6
US-07-789-915A-2
; Sequence 2, Application US/07789915A
; Patent No. 5212058
; GENERAL INFORMATION:
; APPLICANT: Baker, Richard J.
; APPLICANT: Tobias, John W.
; APPLICANT: Varsavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brock, Smith & Reynolds, P.C.
; STREET: Two Millenia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent'n Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,915A
; FILING DATE: 09/11/08
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: XIT-5091AA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1121 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-789-915A-2

Query Match 5.2%; Score 134; DB 1; Length 1121;
Best Local Similarity 24.8%; Pred. No. 0.0001;
Matches 38; Conservative 26; Mismatches 47; Indels 30; Gaps 3;

QY 1 MIVFVFNSSYGFPVEVDSCTSIQLKEVVAKGVADQRLVIFAGKELPHLTVNCD 60
DB 1 MOIFVKTLTGKTTLEVESSTIDNVKSKIQCKEGFPDQQLIFAGKQLEDGRTLSQYN 60
QY 61 LEQGSIVHIVQPRRSRSHETNAGSGDPOSTSEGSIMESRSILTRVDLSHSLTVDSVGLA 120
DB 61 IQKESLHLVLR-----GGDSEED-----EHT----- 85

```



```

QY 104 RVD-LSSHLPVDSVGLAVILDTDSKRDSEAR 135
DB 121 RNRLLAHP-----PFASWENSEAR 141

RESULT 7
US-08-005-002C-2
; Sequence 2, Application US/08C05002C
; Patent No. 5494818
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: Maine
; COUNTRY: U.S.A.
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent: Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/005,002C
; FILING DATE: 15-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,915
; FILING DATE: 08-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: NIT-509-AAZ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 207-363-0558
; TELEFAX: 207-363-0528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1121 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-005-002C-2

Query Match 5.28; Score 134; DB 1; Length 1121;
Best Local Similarity 24.8%; Pred. No. 0.0001;
Matches 38; Conservative 38; Mismatches 47; Indels 30; Gaps 3;

QY 1 MIVFVFNSSYGFVEVDSSTSIQLKXEVAKQGVADQLRVIFAGKELPNHLTVQNC 60
DB 1 MCIEVKLTIGTKTILEVSSDTIDNVKSKIQXEGIPDQCRLIFAGKLEDCRTISDN 60
QY 61 LEQOSIVHIVORRRRSHEHNA-----SGDEPOSTSEGSIWESPSLT 103
DB 61 IQKESLHLVLRGGMHSGCAMLVPSLVKRRKTLAPNTQTASPRALADSLMQLARQVS 120

QY 104 RVD-LSSHLPVDSVGLAVILDTDSKRDSEAR 135
DB 121 RNRLLAHP-----PFASWENSEAR 141

RESULT 8
US-08-487-203A-2
; Sequence 2, Application US/08487203A
; Patent No. 5683904
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.

```

```

; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: Maine
; COUNTRY: U.S.A.
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent: Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,203A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/035,002
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: NIT-509-AJZ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 207-363-0558
; TELEFAX: 207-363-0528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1121 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-203A-2

Query Match 5.28; Score 134; DB 1; Length 1121;
Best Local Similarity 24.8%; Pred. No. 0.0001;
Matches 38; Conservative 38; Mismatches 47; Indels 30; Gaps 3;

QY 1 MIVFVFNSSYGFVEVDSSTSIQLKXEVAKQGVADQLRVIFAGKELPNHLTVQNC 60
DB 1 MCIEVKLTIGTKTILEVSSDTIDNVKSKIQXEGIPDQCRLIFAGKLEDCRTISDN 60
QY 61 LEQOSIVHIVORRRRSHEHNA-----SGDEPOSTSEGSIWESPSLT 103
DB 61 IQKESLHLVLRGGMHSGCAMLVPSLVKRRKTLAPNTQTASPRALADSLMQLARQVS 120

QY 104 RVD-LSSHLPVDSVGLAVILDTDSKRDSEAR 135
DB 121 RNRLLAHP-----PFASWENSEAR 141

RESULT 9
US-09-051-725-25
; Sequence 25, Application US/09091725
; Patent No. 6325141
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Improved methods for transforming Phaffia
; TITLE OF INVENTION: and recombinant DNA for use therein
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster llp
; STREET: 200C Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20006-1588
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO);
; CURRENT APPLICATION NUMBER: US/09/391,725
; APPLICATION NUMBER: US/09/391,725
; FILING DATE: 23-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95203620.0
; FILING DATE: 22-DEC-1995
; APPLICATION NUMBER: EP 96200943.7
; FILING DATE: 11-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Victor Donahue
; REGISTRATION NUMBER: 35,492
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-091-725-25

Query Match 5.0%; Score 123.5; DB 4; Length 158;
Best Local Similarity 24.7%; Pred. NO. 1.6e-05;
Matches 46; Conservative 34; Mismatches 65; Indels 41; Gaps 5;

QY 1 MIVFVFNSSYGPPVEVDSTSLQKKEVAKQGVPAQQLRVIPAGKELPNHILTVQNC 60
Db 1 MQIFVKLTQKTTITLEVSSDTIDNVKAKIQKQKGIPTQCGRLIPAGKELPNHILTVQNC 60
QY 61 LEQOSIVHIVCRPRRRSHETNAGSGDPSSTSGSINWSRLTKRVOLSSHTLPVDSVGLA 120
Db 61 LQKESLTLVLRLR-----GGAKKKKKQVTTTKIKKKKKVKKVAILKY----- 105
QY 121 VILDTSKRSEARGPVKPTNSFLYCKGCHKVCPG-----KLRVCCCTCKQAT 172
Db 106 -----YKVDSD--GKIK-----RLRRCQPOGAGIFMAPHNROTCKG-GLT 147
QY 173 LTLAQG 178
Db 148 YTFNEG 153

RESULT 10
US-08-557-309B-4:
; Sequence 41, Application US/08557309B
; Patent No. 5916572
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodges, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,356
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 21021.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-834-356-41

Query Match 5.0%; Score 128; DB 3; Length 128;
Best Local Similarity 24.8%; Pred. NO. 1.6e-05;
Matches 31; Conservative 28; Mismatches 37; Indels 10; Gaps 2;

QY 1 MIVFVFNSSYGPPVEVDSTSLQKKEVAKQGVPAQQLRVIPAGKELPNHILTVQNC 60
Db 1 MQIFVKLTQKTTITLEVSSDTIDNVKAKIQKQKGIPTQCGRLIPAGKELPNHILTVQNC 60
QY 61 LEQOSIVHIVCRPRRRSHETNAGSGDPSSTSGSINWSRLTKRVOLSSHTLPVDSVGLA 120
Db 61 LQKESLTLVLRLR-----GGVNEPTLEALAKKYNKKKKVKKR 98

RESULT 11
US-08-834-356-41
; Sequence 41, Application US/08834356
; Patent No. 6054115
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodges, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,356
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 21021.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-834-356-41

Query Match 5.0%; Score 128; DB 3; Length 128;
Best Local Similarity 24.8%; Pred. NO. 1.6e-05;
Matches 31; Conservative 28; Mismatches 37; Indels 10; Gaps 2;

QY 1 MIVFVFNSSYGPPVEVDSTSLQKKEVAKQGVPAQQLRVIPAGKELPNHILTVQNC 60

```

```

DB 1 MOIFVKTLTGKTIALVESSDTIENVKAKIQKSG:PPQQRLIFAGKQLEDGRTLDYN 60
QY 61 LEQQSVIHVQRPRLRSHTNASGGDEP--QSTSEGS:WESRSLTR 104
DB 61 IQKESTLHLVLR-----GGWNEPTLEALKKNYKWKVCR 98

RESULT 12
US-08-993-674A-41
; Sequence 41, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David C.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-993-674A-41

Query Match 5.0%; Score 129; DB 3; Length 128;
Best Local Similarity 29.2%; Pred. No. 1.6e-05;
Matches 31; Conservative 28; Mismatches 37; Indels 10; Gaps 2;

QY 1 MIVFVFNSSYGFPEVDSITSLQKEVAKRGQVPADQLRVIFAGKELNHLTVQNC 60
DB 1 MOIFVKTLTGKTIALVESSDTIENVKAKIQKSG:PPQQRLIFAGKQLEDGRTLDYN 60
QY 61 LEQQSVIHVQRPRLRSHTNASGGDEP--QSTSEGS:WESRSLTR 104
DB 61 IQKESTLHLVLR-----GGWNEPTLEALKKNYKWKVCR 98

RESULT 13
US-09-256-976-41
; Sequence 41, Application US/09256976
; Patent No. 6419933
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David C.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-256-976-41

Query Match 5.0%; Score 129; DB 3; Length 128;
Best Local Similarity 29.2%; Pred. No. 1.6e-05;
Matches 31; Conservative 28; Mismatches 37; Indels 10; Gaps 2;

QY 1 MIVFVFNSSYGFPEVDSITSLQKEVAKRGQVPADQLRVIFAGKELNHLTVQNC 60
DB 1 MOIFVKTLTGKTIALVESSDTIENVKAKIQKSG:PPQQRLIFAGKQLEDGRTLDYN 60
QY 61 LEQQSVIHVQRPRLRSHTNASGGDEP--QSTSEGS:WESRSLTR 104
DB 61 IQKESTLHLVLR-----GGWNEPTLEALKKNYKWKVCR 98

RESULT 14
US-09-482-611B-99
; Sequence 99, Application US/09482611B
; Patent No. 6448391
; GENERAL INFORMATION:
; APPLICANT: Garbarino, Joan
; APPLICANT: Belknap, William
; TITLE OF INVENTION: Ubiquitin-Lytic Peptide Fusion Gene Constructs, Protein Products
; TITLE OF INVENTION: Thereafter, and Methods of Making and Using Same
; FILE REFERENCE: 2093-149
; CURRENT APPLICATION NUMBER: US/09/482,611B
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 08/801,025
; PRIOR FILING DATE: 1997-02-19
; PRIOR APPLICATION NUMBER: US 08/279,472
; PRIOR FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 99
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ubiquitin-Lytic Peptide
US-09-482-611B-99

Query Match 5.0%; Score 127.5; DB 4; Length 101;
Best Local Similarity 28.7%; Pred. No. 2.3e-05;
Matches 33; Conservative 26; Mismatches 41; Indels 15; Gaps 2;

QY 1 MIVFVFNSSYGFPEVDSITSLQKEVAKRGQVPADQLRVIFAGKELNHLTVQNC 60
DB 1 MOIFVKTLTGKTIALVESSDTIENVKAKIQKSG:PPQQRLIFAGKQLEDGRTLDYN 60
QY 61 LEQQSVIHVQRPRLRSHTNASGGDEP--QSTSEGS:WESRSLTR 104
DB 61 IQKESTLHLVLR-----GGWNEPTLEALKKNYKWKVCR 98

RESULT 15
US-09-482-611B-101
; Sequence 101, Application US/09482611B
; Patent No. 6448391
; GENERAL INFORMATION:
; APPLICANT: Garbarino, Joan
; APPLICANT: Belknap, William
; TITLE OF INVENTION: Ubiquitin-Lytic Peptide Fusion Gene Constructs, Protein Products
; TITLE OF INVENTION: Thereafter, and Methods of Making and Using Same
; FILE REFERENCE: 2093-149
; CURRENT APPLICATION NUMBER: US/09/482,611B
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 08/801,025
; PRIOR FILING DATE: 1997-02-19
; PRIOR APPLICATION NUMBER: US 08/279,472
; PRIOR FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 99
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ubiquitin-Lytic Peptide
US-09-482-611B-101

```

```

; APPLICANT: Belknap, William
; TITLE OF INVENTION: Ubiquitin-Lytic Peptide Fusion Gene Constructs, Protein Products
; TITLE OF INVENTION: Therefrom, and Methods of Making and Using Same
; FILE REFERENCE: 2093-143
; CURRENT APPLICATION NUMBER: US/09/482,611B
; CURRENT FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 08/801,026
; PRIOR FILING DATE: 1997-02-19
; PRIOR APPLICATION NUMBER: US 06/279,472
; PRIOR FILING DATE: 1994-07-22
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patent version 3.1
; SEQ ID NO 101
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ubiquitin-Lytic Peptide
JS-09-482-611B-101

```

```

Query Match      5.0%   Score 127.5; DS 4; Length 101;
Best Local Similarity 28.7%; Pred. No. 1.3e-05;
Matches 33; Conservative 26; Mismatches 41; Indels 15; Gaps 2;

QY      : MIVFERNSSYGPPVEVDSITSILQKEVAKRQGVADQLRVFAGKELPNHETVONCD 60
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       1 MQFVAKTLTKTITLVESSDTIDNVKAKIEDKGIPTDQQLFAGKQLEDGPTLADYN 60

QY      61 LECQSVH-VQPRPRRSHETNASGGDEPQSTSEGSWEKSLTRVDSLSSHLPVC 115
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       61 IQKSTLHLVLR-R-----GGGSVAKRV-----GRKKKLDK-DRLGVD 100

```

Search completed: October 21, 2003, 21:32:19
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 21:31:10 ; Search time 70 Seconds
(without alignments)
110.025 Million cell updates/sec

Title: US-09-830-703-4
Perfect score: 2570
Sequence: 1 MIVVFBNSSYGPVVEVSD.....CWNCCGEMNRACGDMEDV 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 16746630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA:

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/PCTJS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	28.9	156	US-09-785-548-4	Sequence 4, Appli
2	526	20.5	386	US-10-239-249-2	Sequence 2, Appli
3	242	9.4	46	US-09-864-761-36750	Sequence 16750, A
4	223.5	8.7	474	US-09-808-387-42	Sequence 42, Appl
5	191.5	7.5	753	US-10-146-473-44	Sequence 44, Appl
6	184.5	7.2	838	US-10-149-407-4	Sequence 4, Appl
7	184.5	7.2	838	US-10-149-407-4	Sequence 4, Appl
8	182	7.2	473	US-09-764-864-532	Sequence 822, App
9	182	7.2	617	US-09-764-864-1263	Sequence 1263, Ap
10	182	7.1	617	US-09-850-670-82	Sequence 82, Appl
11	182	7.1	617	US-10-106-698-4649	Sequence 4649, Ap
12	179	7.0	376	US-10-106-698-4550	Sequence 4550, Ap
13	169.5	6.6	433	US-09-764-864-820	Sequence 820, Appl
14	158.5	6.2	403	US-09-996-194-6	Sequence 6, Appl
15	146	5.7	336	US-09-987-107-44	Sequence 44, Appl

16	137.5	5.4	337	10	US-09-987-107-46	Sequence 46, Appl
17	133.5	5.2	154	15	US-10-128-714-8049	Sequence 8049, Ap
18	132	5.1	1548	15	US-10-180-903-2	Sequence 2, Appli
19	131	5.1	421	15	US-10-296-773-7	Sequence 7, Appli
20	128	5.0	522	15	US-10-296-773-2	Sequence 2, Appli
21	127.5	5.0	225	12	US-10-338-421-7	Sequence 7, Appli
22	126.5	4.9	233	12	US-10-311-491-10	Sequence 10, Appl
23	126	4.9	151	12	US-10-332-585-7623	Sequence 7623, Ap
24	124	4.8	75	11	US-09-884-456-35	Sequence 35, Appl
25	124	4.8	75	16	US-10-232-643-3	Sequence 3, Appli
26	124	4.8	151	12	US-10-360-053-10	Sequence 10, App
27	124	4.8	159	12	US-10-360-053-12	Sequence 12, Appl
28	124	4.8	229	15	US-10-137-036-80	Sequence 80, Appl
29	124	4.8	305	9	US-09-866-153-6	Sequence 6, Appli
30	124	4.8	191	9	US-09-866-153-9	Sequence 9, Appli
31	124	4.8	499	10	US-09-887-576-600	Sequence 600, App
32	123.5	4.8	1248	15	US-10-152-724A-4	Sequence 4, Appli
33	123	4.8	156	15	US-10-137-036-67	Sequence 67, Appl
34	123	4.8	504	11	US-09-938-864-410	Sequence 410, App
35	123	4.8	504	15	US-10-125-633A-410	Sequence 410, App
36	123	4.8	504	15	US-10-302-633-410	Sequence 410, App
37	122.5	4.8	2813	10	US-09-886-303-2	Sequence 2, Appli
38	121.5	4.7	1132	15	US-10-189-971-18	Sequence 18, App
39	121.5	4.7	1477	15	US-10-189-971-120	Sequence 20, App
40	121.5	4.7	1477	15	US-10-189-971-9	Sequence 9, Appli
41	121.5	4.7	1593	15	US-10-189-971-14	Sequence 14, Appl
42	121.5	4.7	1593	15	US-10-189-971-4	Sequence 4, Appli
43	121	4.7	76	9	US-09-804-966-16	Sequence 16, Appl
44	121	4.7	76	9	US-09-826-312-13	Sequence 13, Appl
45	121	4.7	76	11	US-09-918-036-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-785-548-4
; Sequence 4, Application US/09755548
; Patent No. US2002015577A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PA
; FILE REFERENCE: SI00005
; CURRENT APPLICATION NUMBER: US/09/785.548
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-548-4

Query Match: 28.9%; Score 744; DB 10; Length 156;
Best Local Similarity 87.2%; Pred. No. 1.4e-61;
Matches 129; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
QY 142 YNSFIYKGPCHKVQPKLRVQGTQCKATLTACQPSQWEDVLIENRSGEQSPCP 201
DB 9 YNSFYVYKGPQQRVQPKLRVQGTQCKATLTACQPSQWEDVLIENRSGEQSPCHP 68
QY 202 GTRAPFFPKCAHPTSDKQTSVALNDITSNERS:PCIACTDVSRLVDFVFCNHRHVCLD 261
DB 69 GTSAPFFPKCAHPTSDKQTSVALNDITSNERS:PCIACTDVSRLVDFVFCNHRHVCLD 128
QY 262 GPHLYCVTRLNDQRFVHDAQIGYSLPCV 280
DB 129 GPHLYCVTRLNDQRFVHDAQIGYSLPCV 156

RESULT 2
US-10-239-249-2
; Sequence 2, Application US/10239249

RESULT 7
 US-10-149-407-4
 ; Sequence 4, Application US/10149407
 ; Publication No. US20030187199A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SOBEU, Gen
 ; APPLICANT: DOUYU, Manabu
 ; TITLE OF INVENTION: No. US20030187199A1e1 Proteins and Genes
 ; FILE REFERENCE: P22483
 ; CURRENT APPLICATION NUMBER: US/10/149,407
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: PCT/JP00/08793
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 838
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-149-407-4

Query Match 7.2%; Score 184.5; DB 12; Length 838;
 Best Local Similarity 24.2%; Pred. No. 3.1e-08;
 Matches 75; Conservative 28; Mismatches 80; Indels 127; Gaps 19;
 QY 217 SDKDTSAVNLITSNRRS-----IPCIAQ---TDVRSPLVFOCNHHRVILDCPF 263
 DB 103 TDNSIFSTNTSSDGLTSISKIGDFICPCLLHRSKORFPD-INTCHRS--CYDCL 159
 QY 264 HLY-----CVTRNDQRQVHDQAGYSLPVAGCNSLIKELHHRFLIG 307
 DB 160 RQVLRIEISERNVNSICTERFNP---HDIRLILS-----DDVMEKY----- 201
 QY 308 EEQYTYQQVGAEECVLQWGVCPSPGGAGLL-----PEQCQRVTCSENGLCGG 360
 DB 202 EEFMLRWLVADPC-----RWCPAPDCGYAVAFPGCASC-----KLTC--GREGCG 247
 QY 361 FVFCRDCKEAYHEGD-CDSLEPSGATSGAYRVKRAAQARWEASKETKKIT----- 414
 DB 248 TFCYCHCKQIWHFNQTC-----AAQERAQSRLRT-RSSSISYSQ 209
 QY 415 -----KPCRCNVPIEK--NGGCMHKCPQPCCKLEWCWNC-----GC 450
 DB 290 ESGAAADIKPCRCACAYIIKVNDGSCNHNMT--AVCGCEFCWLCMKKEISDLHYLSPSGC 347
 QY 451 -----EWNK 454
 DB 348 TFWGKKPWSR 357

RESULT 8
 US-09-764-864-802
 ; Sequence 802, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT223
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 802
 ; LENGTH: 473
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-864-802

Query Match 7.1%; Score 182; DB 10; Length 473;
 Best Local Similarity 26.9%; Pred. No. 2.4e-06;
 Matches 70; Conservative 24; Mismatches 70; Indels 56; Gaps 18;

QY 244 RSPVLVFOCNHHRVCLDCPHLY-----CVTRNDQRQVHDQAGYSLP 287
 DB 16 RARLL-SCPHS--CRCLRHV-LRLEISERVPISCECSELRNP---HDIRLLADP 68
 QY 288 CVAGCNSLIKELHHR---RLGEEQYTRYQOYGAEECVLQWGVCPSPGGAGLL--- 341
 DB 69 -----PLMHKYEEMLRRLASDPCRW-----CPAPDCGYAVIAYG 105
 QY 342 ----PEQQRKVTCEGNGLGGFVFCRDCKEAYHEGD-CDSLEPSGATSGAYRV-DKR 395
 DB 106 CASC-----KUTCERE---SCQTEFCYCHCKQIWHFNQTCVARSQ---QRAQTLRVTKR 154
 QY 396 AAEQARWEASKETIKTKTTPCPCRCNVPIEK--NGGCMHKCPQPCCKLEWCWNC----- 448
 DB 155 TSGLSYGOESGPDII---KPCRCACAYIIKVNDGSCNHNMT--AVCGCEFCWLCMKKEIS 208
 QY 449 -----GC-----EWNK 454
 DB 209 DLHYLSPSGCTFWGKKPWSR 228

RESULT 9
 US-09-764-864-1263
 ; Sequence 1263, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT223
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 1263
 ; LENGTH: 617
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-864-1263

Query Match 7.1%; Score 182; DB 10; Length 617;
 Best Local Similarity 26.9%; Pred. No. 2.4e-06;
 Matches 70; Conservative 24; Mismatches 70; Indels 56; Gaps 18;
 QY 244 RSPVLVFOCNHHRVCLDCPHLY-----CVTRNDQRQVHDQAGYSLP 287
 DB 16 RARLL-SCPHS--CRCLRHV-LRLEISERVPISCECSELRNP---HDIRLLADP 68
 QY 288 CVAGCNSLIKELHHR---RLGEEQYTRYQOYGAEECVLQWGVCPSPGGAGLL--- 341
 DB 69 -----PLMHKYEEMLRRLASDPCRW-----CPAPDCGYAVIAYG 105
 QY 342 ----PEQQRKVTCEGNGLGGFVFCRDCKEAYHEGD-CDSLEPSGATSGAYRV-DKR 395
 DB 106 CASC-----KUTCERE---SCQTEFCYCHCKQIWHFNQTCVARSQ---QRAQTLRVTKR 154
 QY 396 AAEQARWEASKETIKTKTTPCPCRCNVPIEK--NGGCMHKCPQPCCKLEWCWNC----- 448
 DB 155 TSGLSYGOESGPDII---KPCRCACAYIIKVNDGSCNHNMT--AVCGCEFCWLCMKKEIS 208
 QY 449 -----GC-----EWNK 454
 DB 209 DLHYLSPSGCTFWGKKPWSR 228

RESULT 10
 US-09-860-670-82
 ; Sequence 82, Application US/09860670
 ; Patent No. US20020165137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA.2771


```
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 82
; LENGTH: 617
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-860-670-52

Query Match:       7.1%; Score 182; DB 10; Length 617;
Best Local Similarity 26.9%; Pred. No. 3.5e-08;
Matches 70; Conservative 24; Mismatches 70; Indels 96; Gaps 18;

QY      244 RSPVLVFCNHRHVICDCHLY-----CVRLNDRQVFHQAGYSLP 287
        |||:|||||
Db      16 RAPLLI--SCPHRS--CRDCRHRYLRLEISERSRVFISCPESERLNP---HDIRLLADP 68

QY      288 CVAGCPSNLIKELHHF---RIIGEEQYTRVOQGAECEYLXGGVLCCBPBGACGLI--- 341
        |||:|||||
Db      69 -----PLMHKYEEMFLRRYILASDPGCRN-----CPAPDCGYAVIAVG 105

QY      342 ----PEOQRKVTCGGNGLCGFVFRCDCKEAYHEGD-CDSLLEPGATSGAYEV-DKR 395
        |||:|||||
Db      106 CASCP-----KLITCERE----GCQTFFCYHCQKIWHNQICDKVARQ---QRACTLVRTKH 154

QY      396 AAEQARWEASEKETIKKTTPPCRCNVPIEK--NGCGMKWKCPQFOCKLEWCWNC----- 448
        |||:|||||
Db      155 TSLSYGSQSGPDJ-----RPCPRCSAYIIKNMGDSGNMTQC--AVCGCFECMLCKKEIS 209

QY      449 -----GC-----ENNR 454
        |||
Db      209 DLHYLSPSGCTFWGKKFMSR 228

RESULT 11
US-10-106-698-4649
; Sequence 4649, Application US/10106698
; Publication No. US20030103690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PAOS51
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 450
; LENGTH: 376
; TYPE: PRP
; ORGANISM: Homo sapiens
; NAME/KEY: MISC_FEATURE
; LOCATION: 1364;
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: 1370;
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: 1372;
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: 1374;
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4550

Query Match:       7.0%; Score 179; DB 15; Length 376;
Best Local Similarity 26.7%; Pred. No. 3.4e-08;
Matches 48; Conservative 25; Mismatches 57; Indels 50; Gaps 9;

QY      305 ILGEQYTRYQQYGABECV-QMSGV-LCPREGGAGALLPEGQGRKYVCBGGN----- 355
        |||:|||||
Db      12 VYSKWOKRYLOFDIKAFVENNAIKWCPFGCORAV-----ALTKGNTSSGDTLS 64

QY      356 -----GLGCGFVFCROC-KDAYHEGDCDL-----LEPSG--ASTCAIVDYK 394
        |||:|||||
Db      65 FPFLRAPADVCGKGHLFCWCELGAHPGCCQTWKWLQLKITEPKFEELVGVSEAY----- 120

QY      395 AAEGCARWEASEKETIKKTTPPCRCNVPIEKANGGMWMKCQFOCKLEMCANCCEKNR 454
        |||:|||||
Db      121 EDANCLW-----LI-TNSKFUNCKNSPIQKRGGRHNQC--AKCYDFCNILEMWK 171

RESULT 13
US-09-764-864-820
; Sequence 820, Application US/09764864
; Patent No. US20020132793A1
; GENERAL INFORMATION:
; APPLICANT: Roset et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patent'n Ver. 2.0
; SEQ ID NO 820
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-820

Query Match
Best Local Similarity 27.5%; Score 169.5; DB 10; Length 431;
Matches 56; Conservative 24; Mismatches 73; Indels 51; Gaps 13;

QY 266 YCVTRNDROFYDQAOLGYSIFC-VASCP-----NSLIKELHFRILGEEQYTRYQO-----315
DB 14 YLTTRI-----EQNLVNLCTCPADCPAQTGAFIRAI-----VSSPEVISKEYKALLR 62
QY 317 YGAECEVLQMGVLCPPGCGAGLLPECCGQKVTCEGNGCLG-----CGFVFCRCK-- 368
DB 63 GY-VESC---SNLTWCTNP-----CGCERILCRQGLCGGTCSKCGWASCFNCSP 109
QY 369 EAYHEGDCDLSLEPSGATSQAYRVDKRAAEQARWEEASKETIKKT-KPCRCNKVPIEKNG 428
DB 110 EAHYPASC-----GHSQ-WDDGGYDGMVSVAQSKHLAK--SKCFSCQAFIEKNE 161
QY 429 GCMHMKCPQCKLEWNCWNCGEW 452
DB 162 GCLHWTC--AKCNHGFQWRC.LKSN 183
```

```
RESULT 14
US-C9-996-194-6
; Sequence 6, Application US/09996194
; Patent No. US20020151696A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 84242, 8015, 56304, 52999, and 21999,
; FILE REFERENCE: 35800/246590
; CURRENT APPLICATION NUMBER: US/09/996,194
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/250,348
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,073
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/253,878
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,338
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-996-194-6
```

```
Query Match
Best Local Similarity 30.5%; Score 158.5; DB 10; Length 403;
Matches 47; Conservative 16; Mismatches 44; Indels 47; Gaps 11;

QY 331 CRRPGCGAGLL-----PEQQRKVTCEGNGLCCGFVFCRDCCKEAYHEGD-CDSLLEP 382
DB 22 CRRPGCGAVIAYGASCP-----KTCERE---GCCTEFCYHCKQIWHFNQTCDMARQ- 72
QY 383 SGATSQAYRV-DKRAAEQARWEEASKETIKKT-KPCRCNKVPIEK--NGGCMHVKCPQCP 439
DB 73 --QRAOTLVRVTKHTSGSYQGESQPDJ-----KFCRCASAYIIKMGDSGNHWTC--AV 124
QY 440 CKLEWCNC-----GC-----EMNR 454
```

```
DB 125 CGCEFCWLCMKESDCHVILSPSCCTEWGKKPASE 158
RESULT 15
US-C9-987-107-44
; Sequence 44, Application US/09987107
; Patent No. US2002015607A1
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 62/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01662
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patent'n version 3.1
; SEQ ID NO 44
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pT7 H6UEiFp Apo A-1 plasmid
US-C9-987-107-44

Query Match
Best Local Similarity 5.7%; Score 146; DB 10; Length 336;
Matches 53; Conservative 37; Mismatches 64; Indels 56; Gaps 7;

QY 3 VYFRFNSSYGFPEVDSDSILQKEVNAKQVPAQQLRVIFAGKELDNHLYVNCDFLE 62
DB 13 IFVTLTKTITILEVEPSETIENKAKIQCKEGIPPCQRLIFAGKQLDGGATLSQVNIQ 72
QY 63 QOSTIVEVQSPRRSRSHETNAGSCCEPC-----STSEGS: 96
DB 73 KESTLEHLVLRGGSEF--GGCGEPFQGFNFKVVKDLATVYVVLKDSGRDYVVSQPESSA 130
QY 90 -----MEERSLTAVLSSHTLV --- DSVGLAVTLCTSKKDESA 174
DB 121 LGKQNLKLLJCNWDSVTSTPSKREQLGFTVEFNWDEKEFEGLR ---QPMKSDLEEV 186
QY 135 RGPVKPTYNSF-----FVCKGFCCHKVQPGKLVQCG 166
DB 187 KAKVCPYLDEDFQKKWGEEMELIRQ----KVEFLRSLQFG 222

Search completed: October 21, 2003, 21:40:50
Job time : 72 secs
```

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 21, 2003, 21:24:45 ; Search time 43 Seconds
(without alignments)
1037.727 Million cell updates/sec

Title: US-09-830-703-4

Perfect score: 2570

Sequence: 1 MIVVRFNSSYGFPVEVDS.....CWVGCENRACMDHWFV 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	481.5	18.7	357	T23460	hypothetical prote
2	227	8.8	491	F87793	protein C27A12.6
3	224.5	8.7	543	H84724	probable ARI-like
4	218	8.5	644	T03666	hypothetical prote
5	204	7.9	497	G87793	protein C27A12.7
6	202.5	7.9	324	T26983	hypothetical prote
7	201.5	7.8	436	H87793	protein C27A12.8
8	201.5	7.6	514	G84724	probable ARI-like
9	199	7.7	320	T45909	similar to Axiadre
10	191.5	7.5	1753	T03350	hypothetical prote
11	189	7.4	437	T16477	hypothetical prote
12	187.5	7.3	1448	T31653	hypothetical prote
13	187.5	7.1	826	T25555	hypothetical prote
14	181.5	7.1	826	T25555	hypothetical prote
15	179	7.0	488	B86448	hypothetical prote
16	173	6.8	565	F84721	probable RING zinc
17	173	6.7	348	T47494	hypothetical prote
18	172.5	6.7	451	T16481	hypothetical prote
19	167.5	6.5	438	T47498	hypothetical prote
20	162.5	6.3	594	T04783	hypothetical prote
21	162.5	6.3	1209	T52523	hypothetical prote
22	162	6.3	532	T04748	hypothetical prote
23	155	6.0	532	T47487	hypothetical prote
24	150	5.8	222	T47487	hypothetical prote
25	146	5.7	468	A84501	Mutator-like trans
26	144	5.6	464	T48329	probable membrane
27	144	5.6	488	T29562	hypothetical prote
28	143.5	5.6	498	JC5983	protein kinase C-
29	142.5	5.5	154	T46664	ubiquitin/S27a fus

30	141.5	5.5	373	2	F94647	hypothetical prote
31	138	5.4	5282	2	T03454	ALR protein - huma
32	137.5	5.4	158	2	T03590	probable ubiquitin
33	137.5	5.4	518	2	S84536	hypothetical prote
34	137	5.3	377	2	H84578	probable RING zinc
35	135	5.3	546	2	F84647	hypothetical prote
36	134	5.2	80	2	T28305	OPF Msv144 proba
37	132	5.1	1548	2	S34583	setine proteinase
38	131.5	5.1	1548	2	JQ3YR7	ubiquitin / riboso
39	131.5	5.1	618	2	T26611	hypothetical prote
40	130.5	5.1	149	2	JQ3YR7	ubiquitin / riboso
41	130	5.1	77	2	JQ3YR7	ubiquitin - Autogr
42	130	5.1	264	2	S62909	ubiquitin precursor
43	128.5	5.0	839	2	A30126	polyubiquitin
44	129	5.0	154	2	JQ3YR7	ubiquitin / riboso
45	129	5.0	154	2	S55242	polyubiquitin 2 -

ALIGNMENTS

RESULT 1

T23460

hypothetical protein K08E3.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23460

R:McMurray, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19743

A:Accession: T23460

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-357 <WT>

A/Cross-references: EMBL:Z81583; PIDN:CA8C4599.1; GSPDB:GK00021; GESP:K08E3.7

A/Experimental source: clone K08E3

C:Genetics:

A:Gene: GESP:K08E3.7

A:Map position: 3

A:Inserts: 23/3; 72/2; 215/1; 286/2; 311/2

Query Match: 18.7%; Score 481.5; DB 2; Length 357;
Best local similarity 25.6%; Pred. No. 5.1e-10;
Matches 114; Conservative 61; Mismatches 151; Indels 119; Gaps 13;

QY	22	SIQLKEVWAKRGQVPADOLRVIFACKELPHLITVONCLEQOSIVHIVCRERRRSHEN	B1
DB	29	NIEDLKVEKLEIISDELEVVFQKKLSKSTIMRLSLTATQIMLL-RKFNSH-N	B5
QY	52	ASGGDEPQSTSEGSIMESRSLSRVLSLSHTLPVDSVGLAVILTESKRDSZAARGVKPT	141
DB	86	ENGATTAKIITDSSI	100
QY	142	YNFFFYCKGPKGVKGVKRVOCCTCKKATITLAGPSCNVDVILPNKMSRCQSPCF	201
DB	101	LGSEFYWKRN-CGVKRGKLVYCKGTSVLVSEEPQNWSDVLSKSRPAVCESCCT	159
QY	202	GTRAEFFPKCGAHTSOKTSVALNLTNSRSLPGLCTDVASPVNFCQSHHVFOLD	261
DB	160	GLFAEFKFC-----LACNEPAAALTVRGNWCWTECCVCGKEKVFIDLCN-HTCGF	213
QY	262	CFHLYCVTRANDQFVHDACGYSLPCVAGCPNSLAKLHFERILGEEGYTVQVGASE	321
DB	214	CFRV-----RVVCDVHERFHVQCTSYSEVQPKATER	244
QY	322	CV-LQMGVLCRPPGCGAGLL--PEQGGKVTCEGNGLGCGVFVFCQCKEAKHEDCS	378
DB	245	LIAVDKGVTCPNVSCGQSFHFEYDDGKSSCP-----CQFFSPCKC-----FERNVC	295
QY	379	LLPESGATSGAVRVQRAEQARWESAKETIKKTKPKPRCNVFNKNGGWMKRCQCF	438
DB	296	QSE-----DGLRTITLDAITTRCRCHVATERNGGCAHHC--T	332

QY 439 QCKLEWNCOCENWNRACMGDHFWD 463
 DB 333 SCGMDWCFCKTENKEECQNDHWFN 357

RESULT 2

H84724

protein C27A12.6 (imported) - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: F87793

R:Anonymus, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; PMID:9969611; PMID:981916

A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: F87793

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <STO>

A:Cross-references: GB:chr_I; PIDN:AAB91643.1; PID:g2105478; GSPDB:GN00009; CESP:C27A12.

C:Genetics:

A:Gene: C27A12.6

A:Map position: 1

Query Match 8.8%; Score 227; DB 2; Length 491;

Best Local Similarity 26.1%; Pred. No. 5.6e-10;

Matches 74; Conservative 35; Mismatches 94; Indels 80; Gaps 17;

QY 206 BFFPKCGAHTSDXTSVAL---NLITSNRRSIP-----CTACTURSPVLVPCQNRHW 257

DB 94 ERFYE---RP-----DTAFUDACVIRPQCEVIPAGCAECDCISM-DELSG-SCHER-- 143

QY 238 ICLDCHFLYCVTRNDROFVHDAGLGYSLPCVACGPNLSLKE-HHPRILGEEQYTRY--- 314

DB 144 ACACQWQAYLNKI-----VSDAQ--SEIECMA--PNCKL-----LIEDEKVLVSID 197

QY 315 -----QQYGAEECVLQMGVLCPRPGCGALLPEQGO-RKVTCEGNGLGCGF 361

DB 198 PTMSKYRKLMAVSVEINCLRW-----CPGIDCGKAVKVSHPRLVVC-----SCGT 217

QY 362 VFCRCKEAYHEG-DCDSLLEPGSGATQAYRVKRAAEQARWEAEASKETIKKTKPCPRC 420

DB 238 CFCFSCGQWHEPLNCRHL-----KKWIKKQDQSETNWMINANTKCPXC 285

QY 421 NVPIEKNGGCMHMKCPQPCQCKLEWNCWKGCEWNR-----ACMG 458

DB 284 MIPIEKNGGCMRLCTNSGCRYEFCWMCLEPWTXHYQYACNG 326

RESULT 3

H84724

probable AR1-like RING zinc finger protein (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: H84724

R:Lit, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujita, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, C.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; PMID:2009497; PMID:10617197

A:Accession: H84724

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-543 <STO>

A:Cross-references: GB:A5002093; NID:94887759; PIDN:AAC32295.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g31770

A:Map position: 2

Query Match

8.7%; Score 224.5; DB 2; Length 543;

protein C27A12.7 (imported) - Caenorhabditis elegans

RESULT 5

G87753

protein C27A12.7 (imported) - Caenorhabditis elegans

Best Local Similarity 27.2%; Pred. No. 9.7e-10;
 Matches 66; Conservative 32; Mismatches 86; Indels 59; Gaps 14;

QY 234 SIPICTACTD--VASPVLVFCNHRHVICDCHFLYCVTRNGRCQFVHDACLGVSLSF---C 288

DB 125 NIQGGICFESYTRFEEIARVSCGHPY--CATCNAGYITTKIED-----GPGCLRVKCEPESC 178

QY 289 VAGCNSLILKELHHFRILGEEQYTRY--QQYGAEECVLQMGVLCPRPGCGAGLL---PE 343

DB 179 SAAVGKQIMEDVTETKV--NEKYSVILRSYVEDGKKIKW---CSPGCGVAVEFGGSE 232

QY 344 QQRKVTCEGNGC--GGCFVFCRCKEAYHEG-DCDSLLEPGSGATQAYRVKRAAECAWK 402

DB 233 SSSYDVVSCJ-----CSYRFCWNCSEDAHSFVDCQTV-----SKM 266

QY 403 -----PEASKEIKTKTKPCPCNPVPIEKNGCMHMKCPQPCQCKLEWNCWKGCEWNR-- 454

DB 267 IFKQDSEENKNWLANSKPCPECKEPIEKYDCNCNHTCSAP-CGHEFCWIC-KAYRRHS 325

QY 455 -AC 456

DB 326 GAC 329

RESULT 4

T02356

hypothetical protein T9F5.21 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 22-Oct-1999

C:Accession: T02356

R:Vysotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Liu, S.; Li, C.; Ai

Li, Y.; Palm, C.; Shinn, P.; Sun, H.; Davis, R.M.; Ecker, J.R.; Federspiel, K.A.; T

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome 1 BAC T8F5 complete sequence.

A:Reference number: 214666

A:Accession: T02356

A:Status: translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-644 <VYS>

A:Cross-references: EMBL:AC004512; NID:93335331; PID:93335347; GSPDB:GN00059; ATSP:T9F

C:Genetics:

A:Gene: AT9F:T9F5.21

A:Map position: 1

A:Introns: 49/3; 90/2; 125/3; 164/1; 22/3; 254/3; 286/2; 322/2; 352/3; 382/3; 399/3;

Query Match

8.6%; Score 218; DB 2; Length 644;

Best Local Similarity 27.4%; Pred. No. 3.7e-09;

Matches 69; Conservative 32; Mismatches 95; Indels 50; Gaps 14;

QY 215 PTDGNSAVSF-----KQDGGICGFETFSCKLAAACG--HPEDCSWEGYITAIN 171

DB 121 PTDGNSAVSF-----KQDGGICGFETFSCKLAAACG--HPEDCSWEGYITAIN 171

QY 273 PRQFVHDAGLGYSLP---CVAGCFNSLILKELHHFRILGEEQYTRY--QQYGAEECVLQMG 327

DB 172 D----GGCGLTLCRCPDSCRAAVGQDMINLJAPDK--DKQKVTYSFVRSYVEDNRTKW- 224

QY 328 GVLCPREG--GAGLLPEOCORKVTCEGNGC--GGCFVFCRDC-KEAYHEGDCDSLEPSG 384

DB 225 --CPAFGCDYANFVVGSGNYDVNCR-----CCVSFCWNCAEAEHPRVDCITV----- 270

QY 385 ATQAYVDRKAAEQARWEAEASKETIKTKPCPCNPVPIEKNGCMHMKCPQPCQCKLEW 444

DB 271 --SKWVLKNSAESENWY-----LANSKCPCKRPIEKNGCMHITC-TPPCKPEF 320

QY 445 CWNCGCEW 452

DB 321 CWLCLGAW 328

C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001: #text_change 10-May-2001:
 C:Accession: G87793
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MIMD:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_e
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: G87793
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-497 <STO>
 A:Cross-references: GB:chr_I; PIDN:AAB93644.1; PID:G2105479; GSPDB:GN00019; GSEF:C27A12
 C:Genetics:
 A:Gene: C27A12.7
 A:Map position: 1

Query Match 7.9%; Score 204.5; DB 2; Length 497;
 Best Local Similarity 26.3%; Pred. No. 3.5e-08;
 Matches 68; Conservative 34; Mismatches 95; Indels 62; Gaps 16;
 QY 214 HPTSDKDTKVAL---NLITSNRRSIP-----CIATDVRSVVFQCNHRVVICLDCHL 265
 DB 104 HP-----DTIAFLDAQVIRQVEVIPAGDAEDICSM-DELSGLSCNHR--ACAECWQA 154
 QY 266 YCVTRNDRCFVHDAQGLVSLPTVACGNSLI-----KELHFR---ILGEOYTRYQCY 317
 DB 157 YLTNKI-----VSDAQ--SEIEGMA--PNCKLJEDEKVLAKVKTPIAKYRMKNVASY 207
 QY 318 GABECVLQMGVL---CPFGCGAGLLPEQGG-RKVTCEGNGJGCGFVFCRDCKEAYHEG 374
 DB 208 -----IEINALLKMGPGVDCGRTVXVSHGEPRLVCT-----CGSRFCFCGQDMHEP 255
 QY 375 -DCDSLLESGATSOAYRVYDKAAEABEASKETIKTKTKPCRCNVPIEKNGGCMH 433
 DB 256 VNC-----RLKLWYKKNDETSNMINANTKECPKMCAT-ENKGGNCQI 301
 QY 434 KCPQPCCKLEWCMNCGCEW 452
 DB 302 TKNTCKCFQFCWMLGPW 320

RESULT 6
 T15983
 hypothetical protein - curled-leaved tobacco (fragment)
 C:Species: Nicotiana glauca (curled-leaved tobacco)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999: #text_change 20-Sep-1999
 C:Accession: T15983
 R:Borisjuk, N.V.
 submitted to the EMSJ Data Library, October 1996
 A:Reference number: Z18621
 A:Accession: T15983
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-324 <EXL>
 A:Cross-references: EMBL:Y9105; NID:re108749; FID:e280544
 A:Experimental source: somatic embryo
 C:Genetics:
 A:Note: unknown

Query Match 7.9%; Score 202.5; DB 2; Length 324;
 Best Local Similarity 34.6%; Pred. No. 2.9e-08;
 Matches 47; Conservative 14; Mismatches 34; Indels 4; Gaps 3;
 QY 334 PGCGAGLLPE--QGQKVTCEGNGSGCGFVFCRDCK-KEAYHEGDCDSLLESGATSOAY 390
 DB 1 PGCSAVEVDLGGSYDVIC-----CCGFSFCWNCETEARPPVDCVY----- 43
 QY 391 RVDKRAEQAQW-----EASKETIKTKTKPCRCNVPIEKNGGCMHCKCPQPCCKLEW 444
 DB 44 -----AKWLLKNSASENNWVLANSKCPCKRPIEKNGGCMHCTC-TPICKKEP 53

QY 445 CMNCGGCEWNRACMGDH 460
 DB 94 CNLCLGANS-----DH 104

RESULT 7
 H87793
 protein C27A12.8 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May 2001: #text_change 10-May-2001
 C:Accession: H87793
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MIMD:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_e
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: H87793
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-435 <STO>
 A:Cross-references: GB:chr_I; PIDN:AAB93644.1; PID:G2105479; GSPDB:GN00019; GSEF:C27A12
 C:Genetics:
 A:Gene: C27A12.8
 A:Map position: 1

Query Match 7.8%; Score 201.5; DB 2; Length 435;
 Best Local Similarity 25.3%; Pred. No. 4.7e-08;
 Matches 69; Conservative 36; Mismatches 109; Indels 54; Gaps 16;
 QY 206 EPEFKCGAHTSDKTSVAL---NLITSNRRSIP-----CIATDVRSVVFQCNHRV 257
 DB 67 ERYE---RP-----DATTGLDAHVIPRCEELPAGDAECD-CCSL-GEISGLSCNHR-- 116
 QY 258 ICLDCHLYCVRLNDRCFVHDAQGLVSLPTVACGNSLI-----KELHFR---ILGEOYTRYQCY 316
 DB 117 ACTQCKKAVYTKRI-----ANNAQ--SEIEGMAPNCKLJEDEKYNFYTDPTVIATYRK 169
 QY 317 -----YGABECVLQMGVLCPFGCGAGLLPEQGG-RKVTCEGNGJGCGFVFCRDCKEA 370
 DB 170 LIVASVVEINRLKX-----CPGIDCKKAVVSHVHEPRJYTC-----SCGSRFCSCGHD 219
 QY 371 YHEG-DCDSLLESGATSOAYRVYDKAAEABEASKETIKTKTKPCRCNVPIEKNGG 429
 DB 220 WHEPVAC-----RLKLWYKKNDETSNMINANTKECPKMCAT-ENKGGNCQI 301
 QY 430 CMNCKCFCQPCCKLEWCMNCGCEW 452
 DB 266 CMNCKNTACRFECFQWMLGPWP--HGSSWY 296

RESULT 8
 G54724
 Probable APL-like RING zinc finger protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001: #text_change 02-Feb-2001
 C:Accession: G54724
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.F.; Benito, M.L.; Town, C.D.; Fujita, C.Y.;
 M.J. Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, Y.; VarAken, S.E.; Umavathi, L.; Tallon,
 Suss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-769, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MIMD:20083497; PMID:10617197
 A:Accession: G54724
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-514 <STO>
 A:Cross-references: GB:A84420; NID:G887760; PIDN:AAD12296.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2G31760
 A:Map position: 2

Query Match 7.8%; Score 201.5; DB 2; Length 514;

Best Local Similarity 23.98; Pred. No. 5.6e-08;
Matches 60; Conservative 33; Mismatches 82; Indels 77; Gaps 13;

QY 235 IPCACTD--VRSPVLVFOCNHRHVICLDGPHLYCYVTRLNDQFVHDAQLGYSLPCVAGC 292
DB 119 IQGICPFESYTRKEIASVSGHPY--CKTCWTGYITTKIEDG-----PQC 161
QY 293 PNSLIKELHFRILGE-----EYTRY--QQYGAEECVLQMGVLCPRPGCA 338
DB 162 LRVKCPSPSYAVQGMIDDEVTEKKDKKRYFYFLRSYVEDGKKMKW----CPSPGCEC 217
QY 339 GLLPEQCGQRKVTCEGGNGLGCGFVFCRDCKEAYHEG--DCDSLEPSGATQAYRVDKRA 396
DB 218 AVEFGSSGYDVACL-----CSYRFCKNGSEDAHSVDCETV----- 254
QY 397 ABAARK-----EASKETIKTKPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPC 450
DB 255 ---SKWIFKQCESENKMW--LANSKPCPKCRPIEKSHGCHNHVTC--SASGHRFCWICSK 310
QY 451 ENRACVGDHW 461
DB 311 SYS-----CHY 316

RESULT 9
A84725
Similar to Ariadne protein from Drosophila (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 01-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84725
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.; Cowie, C.D.; Fujita, S.Y.;
X.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Niemman, W.C.; White, C.; Eiser, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10611797
A:Accession: A84725
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1542 <STO>
A:Cross-references: GR:AE020093; NID:94887758; PID:AA032294.1; GSPDB:GN00113
C:Genetics:
A:Gene: At2g31780
A:Map position: 2

Query Match 7.88; Score 101; DB 2; Length 542;
Best Local Similarity 25.58; Pred. No. 6.5e-08;
Matches 61; Conservative 31; Mismatches 75; Indels 72; Gaps 13;

QY 235 IPCACTD--VRSPVLVFOCNHRHVICLDGPHLYCYVTRLNDQFVHDAQLGYSLPCVAGC 292
DB 137 IQGICPFESYTRKEIASVSGHPY--CKTCWTGYITTKIEDG-----PQC 179
QY 293 PNSLIKELHFRILGE-----EYTRY--QQYGAEECVLQMGVLCPRPGCA 338
DB 180 LRVKCPSPSYAVQGMIDDEVTEKKDKKRYFYFLRSYVEDGKKMKW----CPSPGCEC 233
QY 339 GLLPEQCGQRKVTCEGGNGLG-----CGFVFCRD--CKEAYHEGDCDSLEPSGATQ 389
DB 234 -----EYVAFPGVNGSSYVSCLSYKFCWNCEDAHSPVDCETV-----SKW 277
QY 390 YRVDKRAAEARNEEASKETIKTKPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPC 448
DB 278 LKXKDESENKMW-----ILAKTPCKPKCRPIEKSHGCHNHVTC--SASGHRFCWICSK 328

RESULT 10
T45209
Hypothetical protein F4P12.390 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45209
B:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Saranoudat, M.

submitted to the Protein Sequence Database, January 2000
A:Reference number: 223016
A:Accession: T45209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <BL0>
A:Cross-references: EMBL:AA332966
A:Experimental source: Cultivar Columbia; BAC clone F4P12
C:Genetics:
A:Map position: 3
A:Introns: 295/2
A:Note: F4P12.390

Query Match 7.78; Score 199; DB 2; Length 320;
Best Local Similarity 27.48; Pred. No. 5.3e-08;
Matches 52; Conservative 37; Mismatches 95; Indels 32; Gaps 10;

QY 229 TSNRRSIPC-ACTVRSPLVFO---CNHRHV--CLOCHELYCYVTRLNDQFVHDAQLGYS 285
DB 107 SSKTATFDCILCVSKS--IESFRIGGCSHPY--CNDVSKYIAKLCQNL-----S 156
QY 286 LPC-VAGCPNSLAKELHFRILGEEQYTRYCYQYGAEECVLQMGVLCPRPGCAAGLPEQ 344
DB 157 IECVSGCGRLSPQQR--QILPKEVFDPMGDALCEAVWMSKKFYCPYKCCSALVLEE 215
QY 345 GCRVTTCEGNGLGCGFVFCRDCKEAYH--EGGDSLEPSGATQAYRVDKRAAEARNE 403
DB 215 SEVXV--KQSECPCHRMVCEVGTQWHEMTCEEFQK--KANEGRDCLJLJATYAK-- 269
QY 404 EASKETIKTKPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPC 448
DB 270 -----QKKWKPCPKCRPIEKSHGCHNHVTC--SASGHRFCWICSK 305

RESULT 11
T00350
Hypothetical protein KIAA0709 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00350
R:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Nomura, H.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98453880; PMID:97374811
A:Accession: T00350
A:Status: preliminary; translated from GS/EMBL/DBEJ
A:Molecule type: mRNA
A:Residues: 1-1753 <ISH>
A:Cross-references: EMBL:AB014609; NID:33327229; PID:BA031683.1; PID:G3327230
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0708

Query Match 7.58; Score 191.5; DB 2; Length 1753;
Best Local Similarity 25.98; Pred. No. 1.3e-06;
Matches 67; Conservative 32; Mismatches 91; Indels 69; Gaps 17;

QY 223 VALNLATSNERSIP----CIACTVRSPLVFOCNH-----RHVICLDGPHLYCYVTR 270
DB 1287 LAGLCVHQVQAVVPRDRCFCV---SPJ---GCDDLSLCCMHYCKSCNVEYLTFR 1340
QY 271 LNDQFVHDAQLGYSLPC-VAGCP-----NSLAKELHFRILGEEQYTRYCYQ-----YGAR 320
DB 1341 I-----ECNLVNLCTCFADCPACTGAF-PAI---VSSPEVSKYKALKLROV-VE 1388

QY 321 ECVLQMGVLCPRPGCGAGLLPEQCGKRVTCGGNGLG-----CGFVFCRDCKEAYHEGDCDSLEPSGATQ 389
DB 1389 SC---SNLTWCTNP-----CQCDRLCGRGCGCTGCKCGWASGFCNCFPEAHYF 1436
QY 374 GDCDSLEPSGATQAYRVDKRAAEARNEEASKETIKTKPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPC 448
DB 1437 ASC-----GMSQ--WVECGGYVDGYSVFAASKHLAKLISKPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPCPC 1489


```

Qy 426 KNGGCMHMKCPQPCKLEWCNC 448
Db 982 KREGCNHQC---GGTHYCWC 1001

RESULT 15
T25555
Hypothetical protein Cl7H11.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25555
R:Johnson, D.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid Cl7H11.
A:Reference number: Z20049
A:Accession: T25555
A:Status: preliminary; translated from GE/EMBL/CDJ
A:Molecule type: DNA
A:Residues: 1-816 <JCH>
A:Cross-references: EMBL:J08047; PDB:1AB37987.1; GSPDB:GN00028; CESP:Cl7H11.6
A:Experimental source: strain Bristol N2; clone Cl7H11.
C:Genetics:
A:Gene: CESP:Cl7H11.6
A:Map position: X
A:Introns: 47/3; 85/3; 122/3; 150/2; 187/2; 260/1; 331/1; 399/1; 448/1; 497/1; 534/3; 59

Query Match 7.1%; Score 181.5; DB 2; Length 816;
Best Local Similarity 23.4%; Pred. NC: 3.3e-06;
Matches 78; Conservative 32; Mismatches 93; Indels 13; Gaps 20;

Qy 157 QPGKLRVCCGTCQATLTLAQQPSCHDVLIPNMSGECQSPDCPOTRAEFFFFKGAHFT 216
Db 84 QPGK-----GKYKECP-CAAKMFGS----APP-KLKG-CQHRSCRACLRQY----- 123
Qy 217 SKDSTVALNLITSNRSIPCIACITDVRSPVLVFCQNHPRHVICJGCPHYCVTRINDRQF 276
Db 124 -----VELS-ITENRVEVPCECSSYLHP-----ND--- 149
Qy 277 VHDALGYSLPVAGCNSITKELHHP---RILGEQYTRYQCYGAECVLQMGVLPQR 333
Db 149 -----IKNLIGDIPITIEKYEAFSLRRYLMTADAK-----CPA 183
Qy 334 PGGAGLLPEQ--GQRKVTCEGNGLCGGVFCDCKEAYHEGD-CDSLPEPSGATSQAY 390
Db 184 PDGQVFATKCAACFOLKQRPD---CGTLFCYHCKREKHSNQ*CD-----EAR 230
Qy 391 RVDKRAEQARWEASK-----ETIKTKTP-----QPRCNVPIEK--NGGCMHMKCPQPQ 439
Db 231 RPEKRSRGLAFBEIYRTGPHQADSGLKPGDVKACPRCKTYIVKMDGDSCHMYC--TM 289
Qy 440 CKLEWCNC-----QC-----EMNR 454
Db 289 CNAEFCWLC-KE-SDLHYLSPTGCTFWGRKPMTR 322

```

Search completed: October 21, 2003, 21:31:53
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: October 21, 2003, 21:22:24 / Search time 24 Seconds
(without alignments;
909.183 Million cell updates/sec

Title: US-09-830-703-4
Perfect score: 2570
Sequence: 1 YIVFVRNKSYPVVDSD.....CWNGCEWNRACMGHFDV 464

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match at
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261.5	10.2	503	1 AR11_DROME	Q94981 Gtscophila
2	237.5	9.2	469	1 AR11_MOUSE	Q921K5 mus musculus
3	236.5	9.2	557	1 AR11_HUMAN	Q9Y4X5 homo sapien
4	227	8.8	492	1 AR12_MOUSE	Q921K6 mus musculus
5	226	8.8	493	1 AR12_HUMAN	Q92376 homo sapien
6	223.5	8.7	474	1 RN14_HUMAN	Q9UBS8 homo sapien
7	214	9.3	485	1 RN14_MOUSE	Q919C0 mus musculus
8	190	7.4	509	1 AR12_DROME	C76924 drosophila
9	189	7.4	551	1 YK27_YEAST	P16113 saccharomyc
10	184.5	7.2	838	1 RN19_HUMAN	Q9NV58 homo sapien
11	184.5	7.2	840	1 RN19_MOUSE	P50636 mus musculus
12	162.5	6.3	292	1 U714_HUMAN	P50876 homo sapien
13	159	6.2	292	1 U714_MOUSE	Q925F3 mus musculus
14	144	5.6	464	1 YMS8_YEAST	Q04638 saccharomyc
15	144	5.6	482	1 AR12_CAEEL	Q22431 caenorhabdi
16	143.5	5.6	498	1 U713_RAT	C62921 rattus norv
17	137	5.3	498	1 U713_MOUSE	Q6W80 mus musculus
18	132	5.1	1877	1 PCXS_MOUSE	Q4592 mus musculus
19	129	5.0	1077	1 SMSA_MOUSE	Q62217 mus musculus
20	128	5.0	76	1 U714_ACACA	P49634 acanthamoeb
21	128	5.0	76	1 UB1Q_NELCR	P13117 neuropept
22	128	5.0	500	1 U713_HUMAN	Q9BY78 homo sapien
23	127	4.9	93	1 UB12_NPVOP	Q05120 Oryza p
24	127	4.9	488	1 U712_HUMAN	Q9W19 homo sapien
25	127	4.9	5376	1 ZAN_MOUSE	C88799 mus musculus
26	126	4.9	75	1 UB1Q_AGJNE	P42740 aglactinam
27	126	4.9	76	1 UB1Q_YEAST	P04838 saccharomyc
28	126	4.9	77	1 UB1L_NPVAC	P16703 atographa
29	125	4.9	76	1 UB1Q_ACECL	P42739 acetabulari
30	125	4.9	76	1 UB1Q_CHURE	P14624 cilianderson
31	125	4.9	76	1 UB1Q_COPCO	P19848 coprinus co
32	125	4.9	76	1 UB1Q_TRYCR	P08565 trypanosoma
33	124	4.8	76	1 UB1Q_ARATH	P59263 arabidopsis

34	124	4.8	76	1 UB1Q_SOYBN	P03993 glycine max
35	124	4.8	76	1 UB1Q_STRPY	P24398 strongyloe
36	123.5	4.8	1700	1 BARI_CHITE	Q03176 citreomus
37	123	4.8	76	1 UB1Q_GHEEL	P14792 caenorhabdi
38	123	4.8	76	1 UB1Q_GECGY	P58669 gesdia cydo
39	123	4.8	76	1 UB1Q_PHYIN	P25899 phytophthor
40	122	4.7	76	1 UB1Q_DICOI	P05618 dictyosteli
41	121	4.7	76	1 UB1Q_DROME	Q92414 drosophila
42	121	4.7	76	1 UB1Q_HUMAN	Q92414 homo sapien
43	121	4.7	76	1 UB1Q_TYVBP	P15174 trypanosoma
44	121	4.7	1598	1 PCXS_BACEL	Q92415 brachyosoma
45	120	4.7	76	1 UB1Q_DEIXA	Q03550 leishmania

ALIGNMENTS

RESULT 1
AR11_DROME STANDARD: PRF: 503 AA.
AC Q94981:
D1 16-OCT-2003 (Ref: 40, Created)
D7 18-OCT-2003 (Ref: 40, Last sequence update)
D8 28-FEB-2003 (Ref: 40, Last annotation update)
DE Ariadne-1 Protein (Ari-1)
GN Ari-1 OR AR1 OR CG5659
CS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID:7227;
RN 1-1
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS.
RX MEDLINE=20341325; PubMed=10890484;
RA Aquilera M., Oliveros M., Martinez-Padron M., Barbas J.A., Ferrus A.;
RI "Ariadne-1: a vital Drosophila gene is required in development and
R: defines a new conserved family of ring-finger proteins."
RL Genetics 155:1231-1244(2000).
RV 1-2
RV SEQUENCE FROM N.A.
RC STEA-N=Serkeley.
RX MEDLINE=20196006; PubMed=1073132;
RA Adams M.C., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanides P.G., Scherer S.E., Li F.W., Hoskins R.A., Gale R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner V., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chapple M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter F.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale C., Bayraktaroglu B., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram R.P., Brandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Broxstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Candia I.,
RA Cherry J.V., Chewley S., Claake C., Davenport J.B., Davies P.,
RA de Paolis B., Delcher A., Deng Z., Deslattes Mays A., Dew I.,
RA Dietz S.N., Dodson R., Eddy M.J., Dowse K., Dugan-Rocha S.,
RA Durkin B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C.,
RA Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S.,
RA Geibart W.M., Glasser K., Glodek A., Gong F., Gottrell J.H., Gu Z.,
RA Guan P., Harris V., Harris N.L., Harvey D., Heiman T.C.,
RA Hernandez J.R., Hickey J., Hostin L., Houston K.A., Howland T.J.,
RA Wei M.-H., Iqbal M., Jain A., Kaul N., Kaul S., Kaul S., Kaul S.,
RA Kennedy S.A., Ketchum K.A., Kim T.-J., Kodira G.D., Kraft C.,
RA Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li C.,
RA Li Z., Liang Y., Lin X., Lin X., Marten H., McIntosh T.C.,
RA McLeod M.P., McPherson D., Merkulov S., Milshina N.V., Mochly N.,
RA Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L.,
RA Muzny D.N., Nelson D.L., Nelson D.R., Nelson X.A., Nixon K.,
RA Nusskern D.R., Pauley C.M., Palazzi M., Pittman G.S., Pat S.,
RA Pollard C., Puri V., Reese V.G., Reiter K., Remington K.,
RA Saunders R.D.C., Scheier F., Shen H., Shue B.C., Siden-Kiers I.,
RA Simpson M., Skupski M.P., Smith T., Spier B., Spradling A.C.,

RA Spapleton M., Strong R., Sun E., Swirakos R., Tector C., Turner R.,
 RA Vetter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A.,
 RA Weinstock G.M., Weissbach J., Williams S.M., Woodage T.,
 RA Worley K.C., Wu D., Yang S., Yao Q.A., Ye C., Yeh R.-F., Zaveri J.S.,
 RA Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N.,
 RA Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W.,
 RA Rubin G.M., Venter J.C.,
 R: "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
 CC CONJUGATING ENZYMES, SUCH AS UBCD10/UBE2L3, AND THEN TRANSFERS IT
 CC TO SUBSTRATES.
 CC -1- SUBUNIT: INTERACTS WITH UBCD10. CAN FORM DIMERS.
 CC -1- SUBCELLULAR LOCATION: MAINLY CYTOPLASMIC.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED, WITH PROMINENT LEVELS IN THE
 CC NERVOUS SYSTEM AND FEMALE GONADS.
 CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS REACHED DURING METAMORPHOSIS AND
 CC MAINTAINED IN THE ADULT.
 CC -1- SIMILARITY: Contains 2 RING-type zinc fingers.
 CC -1- SIMILARITY: Contains 1 IRR-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X98309; CAA66953.1; ..
 CC DR EMBL: X98310; CAA66954.1; ..
 CC DR EMBL: A823507; AA248807.1; ..
 CC DR Flybase: FBgn001744.8; arl1.1.
 CC DR InterPro: IPR002867; Znf.C6HC.
 CC DR InterPro: IPR001841; Znf_Ring.
 CC DR Pfam: PF01485; IRR; 1.
 CC DR SMART: SM00647; IRR; 2.
 CC DR SMART: SM00184; RING; 2.
 CC DR PROSITE: PS00518; ZF_RING_2; FALSE_NEG.
 CC DR PROSITE: PS00089; ZF_RING_2; 1.
 CC DR UBL conjugation pathway; Zinc-finger; Repeat; Coiled coil.
 CC FT DOMAIN 22 27 ASP-RICH (ACIDIC).
 CC FT ZN_FING 133 182 RING-TYPE 1.
 CC FT ZN_FING 203 264 IRR-TYPE 1.
 CC FT ZN_FING 291 335 RING-TYPE 2.
 CC FT ZN_FING 336 361 RING-TYPE 2.
 CC FT DOMAIN 133 201 INTERACTION WITH UBCD10.
 CC FT DOMAIN 341 361 COILED COIL (POTENTIAL).
 CC FT MUTAGEN 150 150 C-x1; IN Aril1-2; LETHAL PHENOTYPE AND
 CC LOSS OF INTERACTION WITH UBCD10.
 CC FT MUTAGEN 309 309 C-x1; IN Aril1-3; LETHAL PHENOTYPE AND NO
 CC LOSS OF INTERACTION WITH UBCD10.
 CC FT SEQUENCE 503 AA; 58932 MW; 5AECE256CF5EC00 CRC64;
 CC
 CC Query Match 10.28; Score 261.5; DB 1; Length 503;
 CC Best Local Similarity 28.88; Pred. No. 2.1e-13;
 CC Matches 79; Conservative 37; Mismatches 97; Indels 61; Gaps 16;
 QY 208 FFKCGAPTSKDTSVNALNLTSSRRSP;PGIACITDVRSP--VLVFOCNHRRVCLDCHEJ 265
 DB 205 FFKC-AHVINFPNATEA;KQKTSRSCQCEICEISQ;PPDSMAGLCGRH--FCNFCWHE 161
 QY 266 YCVTRLNDRQVRDAQLGYS;PCVA-GCPNSJ;KEJHHRFILGEEQY-TRYQYGAEECV 323
 DB 162 YLSTKI-----VARG-LSQITSCAAHGC-DILVDDV-VANLVTDARVRVKYQQLTNSFV 214
 QY 324 LMGSGVIL--CPRPGC-AGLLPEQGRKVTCESGNLGGCGVFCRCKEAYHE----- 373
 DB 215 -ECNGLRWCSFVCTYAVKVPVAPRVRVCK-----CGHVFCAAGCNWHPVACRWJ 267
 QY 374 ----GDDCS-LEPSGA-SQAVRVSKRAEQARWEASKEIKTKTPCPCRCNVPIEKNG 429

Db 268 KXWIKKCD-----DSSTSN-----IAANTKECPRCRVTIKXGG 303
 QY 430 CMHMKCFQPCQKLEWKNCGCEWRNACMSCHWFD 463
 DB 304 CARMVCKNQCKNEFCWVCLGSWEP-HGSSWYN 335
 RESULT 2
 ID ARIL1_MOUSE
 AC Q921K5
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 42, Last annotation update)
 DE Ariadne-1 protein homolog [Ari-1] (Ubiquitin-conjugating enzyme E2-
 DE binding protein 1) (Ubch7-binding protein) (UbcM4-interacting protein
 DE 77) (Fragment).
 GN ARIL1 OR ARI OR UCH7BP OR UTP77.
 OS Mus musculus (Mouse).
 CC Eukaryota; Vertebrata; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [-]
 CC
 CC SEQUENCE OF 1-198 FROM N.A.
 CC STRAIN=CC-1; TISSUE=Embryo;
 CC XLOC=99359765; PubMed=10451819;
 CC Martinez-Noel G., Nienenthal P., Tamura T., Waters K.,
 CC "A family of structurally related RING finger proteins interacts
 CC specifically with the ubiquitin-conjugating enzyme UbcM4."
 CC FEBS Lett. 454:1257-1261(1999).
 CC
 CC SEQUENCE OF 6-469 FROM N.A.
 CC MEDLINE=20341325; PubMed=10850484;
 CC Aguilera M., Olivares M., Martinez-Padron X., Barbas J.A., Ferrus A.,
 CC "Ariadne-1: a vital *Drosophila* gene is required in development and
 CC defines a new conserved family of ring-finger proteins."
 CC Genetics 155:1231-1244(2000).
 CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF THE E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2
 CC UBIQUITIN-CONJUGATING ENZYMES, SUCH AS UBE2L3/UBCVA, AND THEN
 CC TRANSFERS IT TO SUBSTRATES.
 CC -1- SUBUNIT: INTERACTS WITH UBE2L3.
 CC -1- SUBCELLULAR LOCATION: MAINLY CYTOPLASMIC (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- SIMILARITY: Contains 2 RING-type zinc fingers.
 CC -1- SIMILARITY: Contains 1 IRR-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF361001; AA51471.1; ..
 CC EMBL: AJ309777; CAA10275.1; ..
 CC MG: MG11344363; Aril1.
 CC GO: GO:0306511; Ubiquitin-dependent protein catabolism; EPI.
 CC InterPro: IPR002867; Znf.C6HC.
 CC InterPro: IPR001841; Znf_Ring.
 CC Pfam: PF01485; IRR; 1.
 CC SMART: SM00647; IRR; 2.
 CC SMART: SM00184; RING; 2.
 CC PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
 CC PROSITE: PS00089; ZF_RING_2; 1.
 CC UBL conjugation pathway; Zinc-finger; Repeat; Coiled coil.
 CC FT NON_TER 1 1
 CC FT ZN_FING 97 146 RING-TYPE 1.
 CC FT ZN_FING 167 228 IRR-TYPE 1.
 CC FT ZN_FING 255 300 RING-TYPE 2.
 CC FT DOMAIN 344 360 COILED COIL (POTENTIAL).
 CC FT DOMAIN 373 365 INTERACTION WITH UBE2L3.

RX MEDLINE=20341325; PubMed=10890484;
 RI Aquilera K., Oliveros M., Martinez-Padron M., Barbas J.A., Ferrus A.;
 RT "Atrialine-1, a vital Drosophila gene is required in development and
 RL defines a new conserved family of ring-finger proteins."; *Genetics* 155:1231-1244(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=2040251; PubMed=10931946;
 RA Hu R.-X., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
 RA Gu Y.-G., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou C.,
 RA Xu S.-H., Gu J., Jin J.-X., Jin W.-R., Zhang C.-X., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen W.-D., Chen J.-L.,
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 RT axis and full-length cDNA cloning."; *Proc. Natl. Acad. Sci. U.S.A.* 97:9543-9548(2000).
 RJ [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.J., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko B., Madhusiri K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo V.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muziaty S.J.,
 RA Sosak S., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay S.J., Hulyk S.H.,
 RA Vallation D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Raney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakealey R.W., Touchan J.W., Green E.D., Dickson X.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield A.S., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones J.E., Jones J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."; *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903(2002).
 RL [5]
 CC -1- FUNCTION: NIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
 CC CONJUGATING ENZYMES, SUCH AS UBE2J3/UBCX4, AND THEN TRANSFERS IT
 CC TO SUBSTRATES.
 CC -1- SUBUNIT: INTERACTS WITH UBE2J3 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 2 RING-type zinc fingers.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commerce
 CC entities requires a license agreement (See <http://www.isb-sb.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF099149; AAC82469.1;
 CC EXBL: AJ130978; CAAL0276.1;
 CC EXBL: AF181427; AAG09636.1; ALT_FRAME.
 CC EXBL: BC000422; AAH0422.1;
 CC Genew: HGNC:590; ARIH2.
 CC MIM: 605615;
 CC GO: GO:0005634; Cytoc.ueus; TAS.
 CC GO: GO:0008270; Zinc ion binding activity; TAS.
 CC GO: GO:0007275; P.development; TAS.
 CC InterPro: IPR02867; Znf_C6HC.
 CC InterPro: IPR031878; Znf_C6HC.
 CC InterPro: IPR001841; Znf_Ring.
 CC Pfam: PF01495; IRR; 2.

DR SMART: SM00647; IRR; 2.
 DR SMART: SM00184; RING; 2.
 DR SMART: SM00343; ZNF_C2HC; 1.
 DR PROSITE: PS00518; ZF_RING_1; 1.
 DR PROSITE: PS00599; ZF_RING_2; 2.
 KW Cbl conjugation pathway; Nuclear protein; Coiled coil; Zinc-finger;
 KW Repeat.
 FT DOMAIN 4 75 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 22 23 POLY-GLU.
 FT ZN_FING 139 192 RING-TYPE 1.
 FT ZN_FING 209 270 IRR-TYPE.
 FT ZN_FING 302 344 RING-TYPE 2.
 FT DOMAIN 369 400 COILED COIL (POTENTIAL).
 FT DOMAIN 439 492 COILED COIL (POTENTIAL).
 FT CONFLICT 280 281 CA -> JO (IN REF. 2).
 SQ SEQUENCE 493 AA; 578-8 MW; 3CAFFDD327E51013 CRC64;
 Query Match 8.8%; Score 226; EB 1; Length 493;
 Best Local Similarity 26.0%; Pred. No. 1.3e-10;
 Matches 62; Conservative 33; Mismatches 8; Indels 56; Gaps 12;
 QY 237 CIACIV--RSPLVFCNNHRHVICLDCFHLYVTRLDROFVHDAQGLSLPVA--GCP 293
 DV 139 CAVCMQFVKENLSSACIQ--FCRSMECHCSYLKDC-----GVGVGVSCVACQCP 189
 QY 294 NSLIKELHFRFLGEGQVTRVQVGAEECV-----LQXGVLCFRPGGAG--LPEQCOR 347
 DV 190 LRTPEFVFPLPNEBELBEKRYRRLFRDYVESHVQLQ-----LCFGADCPNVRVCEPRAR 245
 QY 348 KYTCEGNSLGGGVFCRDCKEYH--EGDCDSLLSEFSSQATQAVRVKRAAEQARK---- 402
 DV 246 RVQCN-----RQNEVFCEKRCRMVHAPTDCAT-----RKWLTKC 280
 QY 403 --EASKEIKTKTYPCCPCNVPIEKGVGCVHMKCPQPCCKLEKNCNGCCK 452
 DV 281 ADQSETANVISHTKDCPKCNCLIEKGGCGNHMQC--SKRHDFCVMCLGDW 730
 RESULT 6
 RN14 HUMAN
 ID RN14 HUMAN STANDARD; PRT: 474 AA.
 AC QJUBS8; 094793;
 DT 16-OCT-2001 (Rel. 40, Created);
 DT 16-OCT-2001 (Rel. 40, Last sequence update);
 DT 15-SEP-2003 (Rel. 42, Last annotation update);
 DE RING finger protein 14 (Androgen receptor-associated protein 54);
 DE "Ring2 protein" (RFS30) (HRRF52018).
 GN RNF14 (RARA54)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=99055429; PubMed=1020776;
 RA Uski N., Saki N., Yano K., Masuno Y., Saito T., Muramatsu Y.-A.;
 RT "Isolation and characterization of a novel human gene (RPE30) which
 RT encodes a protein with a RING finger motif";
 RL Biochim. Biophys. Acta 1445:232-236(1999).
 RN [2]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC TISSUE=Prostate;
 RX MEDLINE=99195075; PubMed=1008091;
 RA Kang H.-Y., Yeh S., Fujimoto N., Chang C.;
 RT "Cloning and characterization of human prostate coactivator ARA54, a
 RT novel protein that associates with the androgen receptor";
 RL J. Biol. Chem. 274:8570-8576(1999).
 RN [3]
 RP SEQUENCE OF 286-474 FROM N.A. AND SUBCELLULAR LOCATION
 RC TISSUE=Fetal brain;
 PX MEDLINE=99068504; PubMed=9851615;
 RA Uski N., Oda T., Kotaka M., Yano K., Noguchi T., Muramatsu Y.-A.;

RA Wei X.-H., Ibegwan C., Jajali M., Kalush F., Karpen G.H., Ke Z.,
 RA Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kratt C.,
 RA Kravitz S., Kuip D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li C.,
 Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C.,
 RA McLeod M.P., McPherson D., Merkulov G., Milshina N.V., McBarry C.,
 RA Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L.,
 RA Murry D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K.,
 RA Nusskern D.R., Pacleb J.M., Palazcolo M., Pittman G.S., Pan S.,
 RA Pollard J.J., Puri V., Reese Y.G., Reinert K., Remington K.,
 RA Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos T.,
 RA Simpson M., Skupski X.P., Smith T., Spier E., Spradling A.C.,
 RA Stapleton M., Strong R., Sun E., Svitskas R., Tector C., Turner R.,
 RA Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A.,
 RA Weinstein G.M., Weissenbach J., Williams S.M., Woodage T.,
 RA Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yen R.-F., Zaveri J.S.,
 RA Zhan M., Zhang G., Zhao Q., Zhou X., Zheng X.H., Zhong F.N.,
 RA Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W.,
 RA Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 SC Science 287:2185-2195(2000).
 CC - FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
 CC CONJUGATING ENZYMES, SUCH AS UBCD10/UBE2L3, AND THEN TRANSFERS IT
 CC TO SUBSTRATES.
 CC - SUBUNIT: INTERACTS WITH UBCD10.
 CC - SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC - SIMILARITY: Contains 2 RING-type zinc fingers.
 CC - SIMILARITY: Contains 1 IER-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation,
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; AC010169; CAAC9030.1; -
 DR EMBL; AF003456; AAF46821.1; -
 DR FlyBase; FBsm0025186; ast-2.
 DR InterPro; IPR002867; Znf_C6HC.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF01485; IER; 1.
 DR SMART; SM00647; IER; 1.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00589; ZF_RING_2; 1.
 KW Ub1 conjugation pathway; Nuclear protein; Coiled coil; Zinc-finger;
 KW Repeat.
 FT DOMAIN 2 64 ASP/GJU-RICH (ACIDIC).
 FT DOMAIN 123 128 POLY-ALA.
 FT ZN_FING 153 202 RING-TYPE 1.
 FT DOMAIN 222 284 IER-TYPE.
 FT ZN_FING 314 358 RING-TYPE 2.
 FT DOMAIN 397 410 COILED COIL (POTENTIAL).
 FT DOMAIN 452 503 COILED COIL (POTENTIAL).
 SQ SEQUENCE 509 AA; 58627 MW; 9F82E5ACF05E900 CSC64;

Query Match 7.4%; Score 190; DB 1; Length 509;
 Best Local Similarity 25.9%; Pred. No. 1e-07;
 Matches 59; Conservative 33; Mismatches 104; Indels 32; Gaps 9;

QY 229 TSNRSIPCIACITVRSPLVFCQNHRRHVICLDFHLYCVTRLNDRQVHQAQUG-VSLP 287
 DB 145 TPQYRSMQCPVCASSQLQKFKVSLACGHSFKDCWTVYFTQIFQG-----STQIGMAQY 201
 QY 288 CVAGCPNSLIKELHFRILGEEQVTRYQYGAEECVLMQGV-LCPRPCCGAG-LPEQG 345
 DB 202 CNVRVPEDLVLTVPVWRD-----KYQFAKQVYKSHPELRCPGNCGLVCSSEIS 257
 QY 346 QRKVTCEGNGGLCGGVFCRDCKEAYH-EGQCDLSLEPSGATSOAYRYCKEAAQDAWEE 404
 DB 258 AKRAICK-----ACHTGFCRCGMDYHAPTDC-----QVKKWLTKCADDS 298

QY 405 ASKETIKKTTKPCRCACNVPIEKNNGGOMMKCPQPOCKLEWKNCGCEW 452
 DB 299 ETANYISATKDCPKCHTCEIKNGGNHMQC-FNCKHDFQWGLGOW 344

RESULT 9
 ID YKZ7_YEAST STANDARD; PRT; 551 AA.
 AC P36113;
 DT 31-JUN-1994 (Rel. 29, Created)
 DT 31-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 42, Last annotation update)
 DE Hypothetical 63.6 kDa protein in YPS2-GCN3 intergenic region.
 GN YKRL17C.
 OG Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CX NCBI TaxID=4932;
 RN 1;
 PP SEQUENCE FROM N.A.
 RC STRAIN: S288C;
 RA Duestoelcher A., Moestl D., Poehlmann R., Philippsen P.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC - SIMILARITY: Contains 1 RING-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation,
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; Z28242; CA482089.1; -
 DR FIR; S36056; S38386.
 DR SGD; S0001725; YKRL17C.
 DR InterPro; IPR002867; Znf_C6HC.
 DR InterPro; IPR001875; Znf_CCHC.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF01485; IER; 1.
 DR Pfam; PF00597; ZF_C3HC4; 1.
 DR SMART; SM00647; IER; 2.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00589; ZF_RING_2; 1.
 KW Hypothetical protein; Zinc-finger.
 FT ZN_FING 159 224 RING-TYPE.
 SQ SEQUENCE 551 AA; 63569 MW; C04D520D22ED618 CSC64;

Query Match 7.4%; Score 189; DB 1; Length 451;
 Best Local Similarity 24.0%; Pred. No. 1.3e-07;
 Matches 59; Conservative 33; Mismatches 102; Indels 52; Gaps 12;

QY 232 RRSIFCIACITVRSPLVFCQNHRRHVICLDFHLYCVTRLNDRQVHQAQUGVSLP 290
 DB 174 KQDFTCIICDCKKOTETFALECGHEV--CINCYRHY-KDKL-----HEGNIITCYDCSL 225
 QY 291 GCPNSLIKELHFRILGEEQVTRYQYGAEECVLMQ--WGVLCPRPG-----GAGLL 341
 DB 226 ALKNEID-----KVMGHFSSSKLWDSIKSFVCKHNRNKKWCFACKSVLRDTSLL 280
 QY 342 PEOGRK-----VTEGGNGGLCGGVFCRDCR-KAYHEGDCDSLEPSGATSOAYRYCKRA 396
 DB 281 PEYTRLHVSPEVKNSFHR-----PCFNCGFVHSPADC-----KITTAW 320
 QY 397 ASGAKWESAKET-KKTKPCRCNVPIEKNNGGOMMKCPQPOCKLEWKNCGCEW 456
 DB 321 VKARKASEILNWLST-KCEPKCKSVNIEKNNGGOMMKVC--SSCKYECWICGPM--AF 376
 QY 457 GGDHWF 462
 DB 377 HSKNFF 392


```

RESULT 10
RN19 HUMAN
CD RN19 HUMAN STANDARD; PRT: 838 AA.
AC Q9NV58; Q9H5H9; Q9H8V8; Q9UFG3; Q9UFX6; Q9Y4V1;
DT 16-OCT-2002 (Rel. 40, Created);
DT 16-OCT-2002 (Rel. 40, Last sequence update);
DT 28-FEB-2003 (Rel. 41, Last annotation update);
DE RING finger protein 19 (Dorfin) (Double ring-finger protein) (p38
DE protein).
GN RN19
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP SPECIFICITY.
RC TISSUE=Spinal cord;
RX MEDLINE=2113403; PubMed=11237715;
RT "A novel centrosomal ring-finger protein, dorfin, mediates ubiquitin
RT ligase activity."
RL Biochem. Biophys. Res. Commun. 251:706-713(2001).
[2]
RP SEQUENCE OF 142-838 FROM N.A.
RC TISSUE=Hepatoma, Ovarian carcinoma, and Placenta;
RA Isogai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,
RA Masuho Y., Kanemori K., Kawabata A., Hiki T., Kobayashi N.,
RA Inagaki H., Ikema Y., Okamoto S., Okitani R., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Takahashi-Fujii A., Hara H., Tanase T.,
RA Nomura Y., Togiya S., Komai K., Hara R., Takeuchi K., Arita M.,
RA Nabeoka T., Oshima A.;
RT "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 422-824 FROM N.A., AND INTERACTION WITH SP1.
RC TISSUE=Colon;
RX MEDLINE=22431278; PubMed=10976765;
RA Gunther M., Lathier N., Brison C.;
RT "A set of proteins interacting with transcription factor Sp1
RT identified in a two-hybrid screening."
RJ Mol. Cell. Biochem. 210:131-142(2000).
[4]
RP SEQUENCE OF 524-838 FROM N.A.
RC TISSUE=Fetal kidney, and Testis;
RA Bloeker H., Boscher M., Brandt P., Ottenwaelder B., Obermaier B.,
RA Newes H.-W., Gassenhuber J., Wierann S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MEDIATES UBIQUITINATION OF CELLULAR PROTEINS.
CC -!- SUBUNIT: INTERACTS WITH UCH7 AND UCH8. ALSO INTERACTS WITH
CC TRANSCRIPTION FACTOR Sp1.
CC -!- SUBCELLULAR LOCATION: EXPRESSED PRIMARILY IN THE XY BODY OF
CC PACHYTENE SPERMATOCYTES AND IN THE CENTROSOME OF SOMATIC AND GERM
CC CELLS IN ALL PHASES OF THE CELL CYCLE.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED AT LOW LEVELS WITH HIGH
CC LEVELS FOUND IN THE HEART. UBIQUITOUSLY EXPRESSED IN THE CENTRAL
CC NERVOUS SYSTEM.
CC -!- SIMILARITY: Contains 1 IRR-type zinc finger.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC

```

```

DR EMBL: AB023316; BAB33333.1; ALT_FRAME.
DR EMBL: AK021774; BAB1900.1;
DR EMBL: AK027070; BAB15647.1;
DR EMBL: AK021455; BAB14581.1;
DR EMBL: AJ242975; CAB45132.1;
DR EMBL: AJ225096; CAB59284.1;
DR EMBL: AL010251; CAB57700.1;
DR PIR: T34528; T34528.
DR Genbank: HGNC:13432; RNF-9.
DR MIM: 607119;
DR CC: GO:0005813; Centriosome; TAS.
DR CC: GO:0008134; Transcription factor binding activity; TAS.
DR CC: GO:0008270; Zinc ion binding activity; NAS.
DR CC: GO:0002226; Microtubule cytoskeleton organization and biogenesis; TAS.
DR CC: GO:0003464; Protein modification; NAS.
DR InterPro: IPR022867; Zinc finger.
DR InterPro: IPR01841; Zinc ring.
DR Pfam: PF01485; ZRF; 1.
DR Pfam: PF00097; Zinc finger.
DR SMART: SM00647; ZRF; 2.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00519; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS00389; ZF_RING_2; 1.
KW ubi conjugation pathway; Zinc-finger; Metal-binding.
FT ZN_FING 52 117
FT ZN_FING 132 180
FT CONFLICT 645 744
FT CONFLICT 744 744
FT SEQUENCE 838 AA; 93722 MW; 15BD616A52563B6 CRC64;
Query Match 7.2%; Score 184.5; DB 1; Length 938;
Best Local Similarity 24.2%; Pred. No. 4.9e-07;
Matches 75; Conservative 29; Mismatches 80; Indels 127; Gaps 19;
QY 217 SKDTSVALNLTNRSS-----IPCIAC-----TDVSPVLPVQCHRRVILDCDF 263
DB 103 TDKNSTFTSTSSDNGTSTSSKQGDIEPCLCLLRHSKERFPD--INTCHRS--CVDCD 159
QY 264 HLY-----CVTRLDNQFVDAQGYSLPVAGCNLSLKHRRILG 307
DB 160 RCYLRIEISERVNISPECTERNP---HDIFLLS-----DDVLEKY----- 301
QY 308 EQVTRYQYGAEECVIOMGVLPVPGCGASL-----PGGSRKVTCEGNSLGG 360
DB 202 REFMURKLVADPDC-----RWCFAPDCGVAVIAPFCASCP-----KLTC--GREGCG 247
QY 361 FVFCDCCKEAYHEGD--CDSLLPSCGATSCAYSVQKRAEQARMEENSKETINKTT----- 414
DB 248 TFCYCHQIWHPNCTD-----AARQERAGSLFRTIRSSISYSC 259
QY 415 -----KPCPRCNVPIEK--NGGEMHKATCPCKLEKMGNC-----GC 450
DB 290 ESGAAADIKPCPCAAVIVKNGSGSNHMTG--AVGCGEFCGLMKKE--SLHVLSPSGC 347
QY 451 -----ENWR 454
DB 348 TPAGKPKPSR 357
RESULT 11
RN19 MOUSE
ID RN19 MOUSE STANDARD; PRT: 840 AA.
AC P50636; Q9OUJ5;
DT 16-OCT-2002 (Rel. 34, Created);
DT 16-OCT-2002 (Rel. 40, Last sequence update);
DT 16-OCT-2002 (Rel. 40, Last annotation update);
DE RING finger protein 19 (XY body protein) (Xybp) (Garetogenesis
DE expressed protein GEG-154) (UBCM4-interacting protein 100) (Urb100).
GN RNF19 OR XYBP OR GEG-154.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC STRAIN=Swiss; TISSUE=Testis;
RX MEDLINE=20054096; PubMed=10585566;
RA Parra M., del Mazo J.;
RT "Xybp, a novel ring-finger protein, is a component of the XY body of
RT spermatocytes and centrosomes.";
RL Mech. Dev. 90:95-101(2000);
RN [2]
RP SEQUENCE OF 87-245 FROM N.A., AND FUNCTION.
RC STRAIN=CD-1;
RX MEDLINE=99559765; PubMed=10431818;
RA Martine-Noel G., Niedenthal R., Ramura T., Harbers K.;
RT "A family of structurally related RING finger proteins interacts
RT specifically with the ubiquitin-conjugating enzyme UbcM4.";
RL FEBS Lett. 454:257-261(1999);
RN [3]
RP SEQUENCE OF 380-840 FROM N.A., AND DEVELOPMENTAL STAGE.
RC STRAIN=Swiss; TISSUE=Ovary;
RX MEDLINE=95341612; PubMed=7616504;
RA Lopez-Alarcon D.M., Del Mazo J.;
RT "Cloning and characterization of genes expressed during gametogenesis
RT of female and male mice.";
RL J. Reprod. Fertil. 103:323-329(1995);
CC -- FUNCTION: MEDIATES UBILQUITINATION OF CELLULAR PROTEINS.
CC -- SUBUNIT: INTERACTS WITH UBEH7 AND UBEH8. ALSO INTERACTS WITH
CC TRANSCRIPTION FACTOR SP1 (BY SIMILARITY);
CC -- SUBCELLULAR LOCATION: EXPRESSED PRIMARILY IN THE XY BODY OF
CC PACHYTENE SPERMATOCYTES AND IN THE CENTROSOME OF SOMATIC AND GERM
CC CELLS IN ALL PHASES OF THE CELL CYCLE.
CC -- DEVELOPMENTAL STAGE: PREFERENTIALLY EXPRESSED IN BOTH SEXES DURING
CC GAMETOGENESIS.
CC -- SIMILARITY: Contains 1 IRR-type zinc finger.
CC -- SIMILARITY: Contains 1 RING-type zinc finger.
CC -- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 778
CC ONWARD DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF120207; AAF18303.1; -
CC DR EMBL; AF120206; AAF18302.1; -
CC DR EMBL; AF371560; AAG37798.1; -
CC DR EMBL; AF360999; AAK5469.1; -
CC DR EMBL; X71642; CAA50643.1; ALT_FRAME.
CC MGD; MG1:1353623; Rnf19.
CC DR InterPro; IPR002867; Znf_C6HC.
CC DR InterPro; IPR001841; Znf_rings.
CC DR Pfam; PF01485; IRR; 1.
CC DR Pfam; PF00097; Zf-C3HC4; 1.
CC DR SMART; SM00647; IRR; 2.
CC DR SMART; SM00184; RING; 2.
CC DR PROSITE; PS00518; ZF_RING_1; FALSE NEG.
CC DR PROSITE; PS00089; ZF_RING_2; 1.
CC KW Ub1 conjugation pathway; Zinc-finger; Metal-binding.
CC ZN_FING 152 117 IRR-TYPE.
CC FT ZN_FING 132 180 RING-TYPE.
CC SQ SEQUENCE 840 AA; 90632 MW; AC117A35849C023A CRC64;
Query Match 7.2%; Score 184.5; DB 1; Length 840;
Best Local Similarity 24.2%; Pred. No. 4.9e-07;
Matches 75; Conservative 28; Mismatches 80; Indels 127; Gaps 13;

```

```

DB 160 ROYLAR-EISESRVNTSCPECTERENP-----HDIRLLLS-----DVLMEKY----- 201
QY 358 EBOVTRYQVYGAESCVLQMGVLCPRGCGAGLL-----EEOGQPKVTCGNGIGCG 360
DB 222 BEFYLRRLWLVDAPDC-----RWCPADPGYAVIAPGASCSP-----KLTC-----GREGCG 247
QY 361 FVFCRCCKEAYHEGD-CDSLLEPSGATSOAYRYVDKRAEQARMEAEASKETTKKT----- 414
DB 246 TEFVCHCKCIWHENCTCD-----AARQERACSLRDLRTIPSSISYSQ 239
QY 415 -----KQPCRCNVPLEK--NGSCVMKCPSPQCKLEACWNC-----GC 450
DB 293 ESGAAADDIKCPGCAAYIKKNDGSCNPMTC--AVGCEFGWJMYKEISDLAYLSPSGC 347
QY 451 -----EMNR 454
DB 348 TFWGKKPKSR 357
RESULT 12
UT14 HUMAN
ID UT14 HUMAN STANDARD: EPT: 292 AA.
AC P30376;
DT 01-OCT-1996 (Rel. 34, Created);
DT 01-OCT-1996 (Rel. 34, Last sequence update);
DT 29-FEB-2003 (Rel. 41, Last annotation update);
DE Ubiquitin conjugating enzyme 7 interacting protein 4 (UbcM4-
DE interacting protein 4) (RING finger protein 141).
GN ERF144 OR UBC97124 OR KIAA0161
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN 1.
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96281124; PubMed=8724849;
RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;
RT Prediction of the coding sequences of unidentified human genes. V.
RT the coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT analysis of cDNA clones from human cell line KG-1.
RL DNA Res. 3:17-24(1996).
CC -- SUBUNIT: Interacts with UBE2L3 (By similarity).
CC -- SUBCELLULAR LOCATION: Nucleus (By similarity).
CC -- SIMILARITY: Contains 1 RING-type zinc finger.
CC -- SIMILARITY: Contains 1 RING-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D79983; BAA11478.1; -
CC DR Genes; HGNC:20457; RNF144.
CC DR InterPro; IPR002867; Znf_C6HC.
CC DR InterPro; IPR001841; Znf_ring.
CC DR Pfam; PF01485; IRR; 1.
CC DR Pfam; PF00097; Zf-C3HC4; 1.
CC DR SMART; SM00647; IRR; 2.
CC DR SMART; SM00184; RING; 2.
CC DR PROSITE; PS00518; ZF_RING_1; 1.
CC DR PROSITE; PS00089; ZF_RING_2; 1.
CC KW Ub1 conjugation pathway; Zinc-finger; Nuclear protein.
CC ZN_FING 20 69 RING-TYPE.
CC FT ZN_FING 20 69
CC SQ SEQUENCE 292 AA; 32859 MW; E2F86A3A3E8514F2 CRC64;
Query Match 4.3%; Score 162.5; DB 1; Length 292;
Best Local Similarity 24.6%; Pred. No. 9.3e-06;
Matches 62; Conservative 36; Mismatches 91; Indels 63; Gaps 15;

```

```

CY 215 PTSDKTSVALNITENRRBIPICACTDVRSPVLVFCQNRHNVICLDPHFYCVTRLNDR 274
DB 2 PTWD-----LALDPLNSCKL---CLGYPVEQMTIAQCC-----CIFC--TLCC-----K 43
CY 275 QFVHDAOLGYSLPCVAGCPNLSIKELHFR-----ILGEEVTVQCVGAE-EGVLQMG 327
DB 49 QYV-ELLIKLEGLETATSCPAAACPKGHLQENIECYVAAEIMQYKLPQREVLFPDPC 107
CY 328 GVLCPRPGCA-----GLPFGQGRKVTCEGNGLGCGFVFCRDCKEAYHEGD-CUSL 379
DB 108 RTWCPASTCQAVCQLOVGS---QTPQVCK-----ACRMFCSTCKSAWHFPGCGCFET 159
CY 380 LEPS---GATSOAYRVDKRAEAOARVEASKETIKTKTPKPCRCQVPEKNGGCMHMKCP 435
DB 160 MPITFJEGETSAAPKEBDA-----PIKPCPKRVYIERDEGCAQMC- 203
CY 437 QPOCKLEWCNC 448
DB 204 -KXCKHAFQVC 214

RESULT 13
UT14 MOUSE STANDARD; PRT: 292 AA.
AC Q225F3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin conjugating enzyme 7 interacting protein 4 'Ubcm4'-
DE Interacting protein 4 (RING finger protein 144).
GN RNFI44 OR UBE7IP4 OR UIP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
EN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Noel G., Pringa E., Harbers K.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
(2)

```

```

SEQUENCE FROM N.A.
RX MEDLINE=22354683; PubMed=12477932;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Berger J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.V., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaner C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleto K., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Teshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lottelino N.A., Peters G., Abramson R.D., Millard S.C.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Vitek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay D.J., Rulyk S.W.,
RA Villalobos D.K., Yuzny D.M., Sodergren E.J., Xu X., Gibbs S.A.,
RA Fahy G., Helton E., Ketteran V., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Wyers R.W.,
RA Butterfield V.S.N., Krzywinski M., Skalska U., Smalusz D.E.,
RA Schnorcher A., Schein C.E., Jones S.C.M., Matra V.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SUBUNIT: Interacts with UBE2E3 by similarity.
CC -!- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copy-right. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/
CC or send an email to license@sib.ch).
CC
DR ENBL; AF362996; AK51468 1;
DR ENBL; AK389886; BAC30058 1;
DR ENBL; BC30187; AAH30187.1;
DR VGD; VGI:134401; Rn1144;
DR InterPro: IPR002867; Znf_C6HC;
DR InterPro: IPR001841; Znf_Ring;
DR Pfam: P501485; IPR: 1;
DR PRCSITE; P500518; ZF_RING_1;
DR PRCSITE; P500589; ZF_RING_2;
KW Ubiquitination pathway; Zinc-finger; Nuclear protein.
SC SEQUENCE 292 AA; 12845 MW; 558ATCBAL52777D CRC64;

Query Match 6.2%; Score 159; DB 1; Length 292;
Best Local Similarity 22.4%; Pred. No. 166-05;
Matches 66; Conservative 36; Mismatches 94; Gaps 15;

CY 173 LTLAQPSCWCDVLIENMSGECSFDCPTAEFFKCGNHPSTDKTSVALNLTSNR 232
DB : MTARVPTWDLALDPLVSCXK:C-----LGEVP-ABQMTTIA----- 36
CY 233 RSPICACTDVRSPVLVFCQNRHNVICLDPHFYCVTRLNDRQVHCAQLGSLPCVAGC 292
DB 37 -CCQICFCT-----LCL-----KQV-ELLIKLEGLETATSC 65
CY 293 PMSLIKEHFR-----LGEQVTRYQVQYAE-EGVLQNGVLCRPGCA----- 336
DB 66 FFAACFKCHLQENIECYVAAEIMQYKLPQREVLFPDPCWCPASTCQAVCQD: 125
CY 339 GLPFGQGRKVTCEGNGLGCGFVFCRDCKEAYHEGD-CDSLPPS---GATSOAYRVDK 394
DB 126 GL---QTPQVCK-----ACMFCSCAKRNHPGCGCFETITFLPGTSAFPMER 177
CY 395 RAEOARVEASKETIKTKTPKPCRCQVPIKNGGCMHMKCPQPOCKLEWCNC 448
DB 178 GDA-----PIKPCPKRVYIERDEGCAQMC--KXCKHAFQVC 214

RESULT 14
YMG8_YEAR:

```

```

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs".
RL Nature 420:563-573 (2002).
RN [3]

```

```

ID YMG8_YEAST STANDARD: PRT: 464 AA.
AC Q04638;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 54.1 kDa protein in DAK1-ORC1 intergenic region.
GN YML068W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Raymond M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XII."
RL Nature 387:90-93(1997).
CC -!- SIMILARITY: TO C.ELEGANS F56D2.2.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z38114; CAA86232.1;
CC F1R: S48329; S48329.
CC GO: GO:0004533; YML068W.
CC InterPro: IPR002867; Znf_C6HC.
CC InterPro: IPR001841; Znf_ring.
CC Pfam: PF01485; IPR: 2.
CC SMART: SM00647; IPR: 2.
CC SMART: SK00184; RING_2.
CC PROSITE: PS00518; ZF_RING_1; FALSE NEG.
CC PROSITE: PS00089; ZF_RING_2;
CC KW Hypothetical protein; Zinc-finger.
FT ZN FING 180 233 RING-TYPE.
SQ SEQUENCE 464 AA; 54095 MW; F48372B742AF63D CRC64;
Query Match 5.6%; Score 144; DB 1; Length 464;
Best Local Similarity 22.6%; Pred. No. 0.0041;
Matches 68; Conservative 34; Mismatches 119; Indels 80; Gaps 26;
QY 210 KCAHPTSD--KTSVALNLTISNRRSPCIACDVRSPLVVFQ-----CNRRHVICLD 262
DB 151 KCUTQDYLKQISEATLQKVSRYHCCICMEXGVAMIKLPENANVHYLCRG 210
QY 263 FHLVYVTRLNDRQF--VHDAQIGYS-----IPCVAGCNSLIKEL-- 300
DB 211 AKSYFTAMICENRISVRCPQCEYKELKLEDEKSYKMKLKALFTPLI---FVSPFKEVID 267
QY 301 ----HHFRLGEEQVRYQVGAEECVLQMGVLCPRGCGAGLLPQ--GQRVTCGG 354
DB 268 TELCERYKMFNQATRLSKVCPAC-----VTCRR--CDSWTKEDLDLDMIQO-- 317
QY 355 NLGCGGVFCRDCKEAYH--EGDCDSLLPSGATQVAYRVD-----KRAAEQARN-- 402
DB 318 ---KCHFVEFCFCLHAWHGNNKGGKVSLSLDTIEEYLDQVTVSYENKXLE-AKYGR 373
QY 403 -----EASKETIK--ITKPCPRNVPIKNGGGMWKKCPQPOCKLEWNC 448
DB 374 VLELVNDYLAEXMLDLAIKKEGSLNQRCFKGVYVSEGCNRYKC--EVCGLFGFIC 431
QY 449 G 449

```

```

DB 412 G 432
RESULT 15
AR12 CAEE,
ID AR12 CAEE, STANDARD: PRT: 482 AA.
AC Q22431;
DT 16-OCT-2001 (Rel. 40, Created)
DT 23-FEB-2003 (Rel. 41, Last sequence update)
DT 23-FEB-2003 (Rel. 41, Last annotation update)
DE Probable ariadne-2 protein (Ari-2).
GN T12E12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Perodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Brisc01.N2;
RA Bradshaw H., Steliys L.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: MIGHT ACT AS AN E3 UBCUITIN-PROTEIN LIGASE, OR AS PART
CC OF E3 COMPLEX, WHICH ACCEPTS UBCUITIN FROM SPECIFIC E2 UBCUITIN-
CC CONUGATING ENZYMES, SUCH AS UBC-2/UBE2L3, AND THEN TRANSFERS IT
CC TO SUBSTRATES (BY SIMILARITY).
CC -!- SUBUNIT: INTERACTS WITH UBC-2 (BY SIMILARITY).
CC -!- SUPRACELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Contains 2 RING-type zinc fingers.
CC -!- SIMILARITY: Contains 1 BR-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U61944; AAB3321.2;
CC WcMmpP: T12E12.1; E31725.
CC InterPro: IPR002867; Znf_C6HC.
CC InterPro: IPR001841; Znf_ring.
CC Pfam: PF01485; IPR: 2.
CC SMART: SM01485; IPR: 2.
CC SMART: SM0647; IPR: 2.
CC PROSITE: PS00518; ZF_RING_1; FALSE NEG.
CC PROSITE: PS00089; ZF_RING_2; FALSE NEG.
CC KW Ub1 conjugation pathway; Nuclear protein; Coiled coil; Zinc finger.
FT ZN FING 2 60 ASP/GLU-RICH /AC-DICD.
FT ZN FING 129 177 RING-TYPE 1.
FT ZN FING 188 260 IPR-TYPE.
FT ZN FING 287 330 RING-TYPE 2.
FT ZN FING 433 459 COILED COIL /POTENTIAL.
SQ SEQUENCE 482 AA; 55808 MW; DFD258B911E0F6F CRC64;
Query Match 5.6%; Score 144; DB 1; Length 482;
Best Local Similarity 21.7%; Pred. No. 0.0043;
Matches 58; Conservative 40; Mismatches 107; Indels 62; Gaps 13;
QY 204 RAETFFKC--GAHPTSDKTSVALNLTISNRRSPCIACDVRSPLVVFQCNRRHVICLD 260
DB 98 RNDFLKCHIDAKPKPKKLSSTQSVLAKGV-----CSVCAMDSYTELPHLTCG--RCFCE 151
QY 261 DCFHLYVYVTRLNDRQFVHDAQIGYSJPCVAG-- -- --CENGLKEHHFRL 306
DB 152 HCKGHVSRLSE-----GVASRIECMESECEVYAPSEFVATIKNSPVIKLYERFL 204
QY 307 GEEQVRYQVGAEECVLQMGVLCPRGCGAGLLPQ--GQRVTCGGNGLGCGVFCRD 366

```

```

Db      205 LRDMVNSHPH--LKFCV---GNECFVIRSTEVKE---KRVTC-----XQCHTSECVX 249
QY      367 CKRAYH-EGDCDSLLEPSGATSCAYRVDKRAAEQARWEEASKEGKTKKPCPCRCNVPIE 425
Db      250 CGADYHAPTSCETI-----KQWTKCADJSETANYISAHTKDCPOCHSCIE 295
QY      425 KNGGCMHMKPOPOCKLEWKNCGGEM 452
Db      296 KAGGCNHQC--TRGRHFCWMCFGDW 320

```

Search completed: October 21, 2003, 21:29:15
 Job time : 26 secs


```

DR SMART; SM00213; UBO; 1.
DR PROSITE; PS00139; THIOLESTERASE_CYS; 1.
DR PROSITE; PS0053; UBIQUITIN_2; 1.
SQ SEQUENCE 464 AA; 5167 AA; 55742859A1B086 CRC64;

Query Match      100.0%; Score 2570; DB 11; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.1e-233;
Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIVFVFNSSYGFPPVEVDSDTSILQKQWAKRGVDPADQLRVIFAGKELPNHLTVQKCD 60
DB 1 MIVFVFNSSYGFPPVEVDSDTSILQKQWAKRGVDPADQLRVIFAGKELPNHLTVQKCD 60
QY 61 LEQOSTVHVIVPRPRSHETNASGGDEPQSTSEGSIMESKSLTRVQLSSHTLPVDSVGLA 120
DB 61 LEQOSTVHVIVPRPRSHETNASGGDEPQSTSEGSIMESKSLTRVQLSSHTLPVDSVGLA 120
QY 121 VILDTSKRDSAAAGP-VKPTNSFFIYCKGPKHVPGKLVCCGCKCAT-TLAQGS 180
DB 121 VILDTSKRDSAAAGP-VKPTNSFFIYCKGPKHVPGKLVCCGCKCAT-TLAQGS 180
QY 181 CMDDVLIPNRMGEGCQSPDCTRAEFPFKCGAHTSKDTSVALNLTITSNRRSIPCIAC 240
DB 181 CMDDVLIPNRMGEGCQSPDCTRAEFPFKCGAHTSKDTSVALNLTITSNRRSIPCIAC 240
QY 241 TVRSPLVLFQCNHRHVLCLOCFHLYCVTRUNDQFVADACLGYSLPCVAGCCPNSLIKEL 300
DB 241 TVRSPLVLFQCNHRHVLCLOCFHLYCVTRUNDQFVADACLGYSLPCVAGCCPNSLIKEL 300
QY 301 HHFRILGEEQTRYQYQYGAEECVLQMGGVLCPRPGCGAGLLPEGGQKVKTCGGNGLGC 360
DB 301 HHFRILGEEQTRYQYQYGAEECVLQMGGVLCPRPGCGAGLLPEGGQKVKTCGGNGLGC 360
QY 361 FVFCRCCKEAYHEGDCSLLEPSGATSOAYRVDKRAEQARWEASKEET-KKTKPCPR 420
DB 361 FVFCRCCKEAYHEGDCSLLEPSGATSOAYRVDKRAEQARWEASKEET-KKTKPCPR 420
QY 421 NVPIEKNGGCMHMKCPQPCCKLEWNCNGCENRACNGCHMFQV 464
DB 421 NVPIEKNGGCMHMKCPQPCCKLEWNCNGCENRACNGCHMFQV 464

RESULT 2
Q9JMK6 PRELIMINARY; PRT; 465 AA.
AC Q9JMK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parkin.
GN PARKIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1.
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Soda Y., Imai Y., Takahashi R.;
RT "Molecular cloning of rat Parkin gene."
RL Submitted (APR-2000) to the EMBL/GenBank/CDDB databases.
RN 2.
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Gu W. J., Abbas N. E.;
RT "Parkin CDNA."
RL Submitted (DEC-1999) to the EMBL/GenBank/CDDB databases.
DR EMBL; AF257234; AAF6666.1;
DR EMBL; AF257234; AAF6666.1;
DR HSSP; P02248; 1UBI.
DR InterPro; IPR003197; Parkin.
DR InterPro; IPR000169; SHPOT_acsite.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR002867; Znf_C6HC.
DR Pfam; PF01485; IIR; 1.
DR Pfam; PF0240; ubi-ubiquitin; 1.
DR PRINTS; PR01473; PARKIN.
DR PRINTS; PR03348; UBIQUITIN.
DR SMART; SM00647; IIR; 2.

```

```

DR Pfam: PF01485; IIR: 1.
DR Pfam: PF02440; ubi:quitin; 1.
DR PRINTS: PR01475; PARKIN.
DR PRINTS: PR03348; UBIQUITIN.
DR SMART: SMC0647; IIR: 2.
DR SMART: SMC0213; JBO: 2.
DR PROSITE: PS00139; THIOLE PROTEASE_CYS; 1.
DR PROSITE: PS00093; UBIQUITIN_2; 1.
SQ SEQUENCE 465 AA; 51308 MW; 513CF17CAD6D942B CRC64;

Query Match
Best Local Similarity 94.6%; Score 2455.5; DB 11; Length 465;
Matches 440; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

CY 1 MIVFVFNSSYGFPVEVSDTSILQKEVAKRGVADQLRVIFAGKELPHLTVQNC 60
DB 1 MIVFVFNSSYGFPVEVSDTSILQKEVAKRGVADQLRVIFAGKELPHLTVQNC 60
CY 61 LEQOSIVHIVORPRRSKSHETNAGGDEPOSTSEGSWESRSITRYDLSHTLPVCSVGLA 120
DB 61 LEQOSIVHIVORPRRSKSHETNAGGDEPOSTSEGSWESRSITRYDLSHTLPVCSVGLA 120
CY 121 VILDTSDKSDSEAAARGP-VKPTVNSFFIYCKGPKHVQKURVQCGTCKQATLTLAGCP 179
DB 121 VILDTSDKSDSEAAARGP-VKPTVNSFFIYCKGPKHVQKURVQCGTCKQATLTLAGCP 179
CY 180 SCWDDVLIPNMSGECSPDCGTFRAEFKCGAHTSDKTSVALNLTNSRSIPGIA 239
DB 181 SCWDDVLIPNMSGECSPDCGTFRAEFKCGAHTSDKTSVALNLTNSRSIPGIA 240
CY 240 CTDRSPVLVFCQNHRRVILCDLHFLYCVTRLNDROFVHDAQLGYSLPVACGPNLSIKE 299
DB 241 CTDRSPVLVFCQNHRRVILCDLHFLYCVTRLNDROFVHDAQLGYSLPVACGPNLSIKE 300
CY 300 LHFRILGEEQYTRYQYGAEEVLMGVLCPRGCGAGLLPEQGGKRVCEGNGJG 359
DB 301 LHFRILGEEQYTRYQYGAEEVLMGVLCPRGCGAGLLPEQGGKRVCEGNGJG 360
CY 360 GFVFCRCDEAYHEGDCDSLEPSGATSOAYRYDKRAAEQAEWESASNETIKKTKPCPR 419
DB 361 GFVFCRCDEAYHEGDCDSLEPSGATSOAYRYDKRAAEQAEWESASNETIKKTKPCPR 420
CY 420 CNTPIKNGGCMHMKCPQPOCKLEWNCGCENWACNACMGDHFVDV 454
DB 421 CNTPIKNGGCMHMKCPQPOCKLEWNCGCENWACNACMGDHFVDV 465

RESULT 4
Q8K5C6 PRELIMINARY; PRT: 489 AA.
AC Q8K5C6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parkin isoform.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1.
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RJ Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF381277; AA021452.1;
DR InterPro: IPR003977; parkin.
DR InterPro: IPR000169; SHP1prot. acsite.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF01485; IIR: 1.
DR PRINTS: PR01475; PARKIN.
DR SMART: SMC0647; IIR: 2.

```

```

DR SMART: SMC0213; UBC: 1.
DR PROSITE: PS00139; THIOLE PROTEASE_CYS; 1.
DR PROSITE: PS00093; UBIQUITIN_2; 1.
SQ SEQUENCE 489 AA; 54412 MW; 49F53B9F355B2A85 CRC64;

Query Match
Best Local Similarity 94.7%; Score 2433.5; DB 11; Length 489;
Matches 440; Conservative 12; Mismatches 12; Indels 2; Gaps 2;

CY 1 MIVFVFNSSYGFPVEVSDTSILQKEVAKRGVADQLRVIFAGKELPHLTVQNC 56
DB 1 MIVFVFNSSYGFPVEVSDTSILQKEVAKRGVADQLRVIFAGKELPHLTVQNC 60
CY 57 ----- CMCLEQOS:VIIIVORPRRSKSHETNAGGDEPOSTSEGS 96
DB 61 DFCCHKSHLAVNLSQQQVTCNCDLEQOS:VH:VORPRRSKSHETNAGGDEPOSTSEGS 120
CY 97 WESRSITRYDLSHTLPVCSVGLAVILDTSDKSDSEAAARGP-VKPTVNSFFIYCKGPKHK 155
DB 121 WESRSITRYDLSHTLPVCSVGLAVILDTSDKSDSEAAARGP-VKPTVNSFFIYCKGPKHK 180
CY 156 VDPGKURVQCGTCKQATLTLAGCPSCWDVLFNMSGECSPDCGTFRAEFKCGAHP 215
DB 181 VDPGKURVQCGTCKQATLTLAGCPSCWDVLFNMSGECSPDCGTFRAEFKCGAHP 240
CY 216 TSDKTSVALNLTNSRSIPGIACTVRSVPLVFCQNHRRVILCDLHFLYCVTRLNDRO 275
DB 241 TSDKTSVALNLTNSRSIPGIACTVRSVPLVFCQNHRRVILCDLHFLYCVTRLNDRO 300
CY 276 FVHDAQLGYSLPVACGPNLSIKEJHFRILGEEQYTRYQYGAEEVLMGVLCPRPG 335
DB 301 FVHDAQLGYSLPVACGPNLSIKEJHFRILGEEQYTRYQYGAEEVLMGVLCPRPG 360
CY 336 CGAGLLPEQGGKRVCEGNGJGCGFVFCRCDEAYHEGDCDSLEPSGATSOAYRYDKR 395
DB 361 CGAGLLPEQGGKRVCEGNGJGCGFVFCRCDEAYHEGDCDSLEPSGATSOAYRYDKR 420
CY 396 AAEQAEWESASNETIKKTKPCPRCNVPIEKNGGCMHMKCPQPOCKLEWNCGCENWRA 455
DB 421 AAEQAEWESASNETIKKTKPCPRCNVPIEKNGGCMHMKCPQPOCKLEWNCGCENWRA 480
CY 456 CMGDHFVDV 464
DB 481 CMGDHFVDV 489

RESULT 5
Q8K5C6 PRELIMINARY; PRT: 459 AA.
AC Q8K5C6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parkin (Fragment)
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1.
RP SEQUENCE FROM N.A.
RC T-STRIDE=Brain;
RX MEDLINE=2015312; PubMed=15686154;
RA D'Agata V., Zhao W., Cavallaro S.;
RT "Cloning and distribution of the rat parkin mRNA.";
RJ Brain Res. Mol. Brain Res. 75:145-149(2000).
DR EMBL: AF168304; AAF34874.1;
DR HSP: P02248; UBI.
DR InterPro: IPR03977; parkin.
DR InterPro: IPR000169; SHP1prot. acsite.
DR InterPro: IPR000626; Ubiquitin.
DR InterPro: IPR002867; Znf_C6HC.
DR Pfam: PF01485; IIR: 1.
DR PRINTS: PR01475; PARKIN.
DR SMART: SMC0213; UBIQUITIN; 1.

```



```

DR PRINTS; PRO1475; PARKIN.
DR PRINTS; PRO0348; UBICUTIN.
DR SMART; SMO0647; IER; 2.
DR SMART; SMO0213; UBQ; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00553; UBICUTIN_2; 1.
DR NON_TER; 459
SQ SEQUENCE 459 AA; 50864 MW; 818BAE7559A93D39 CRC64;

Query Match
Best Local Similarity 93.9%; Score 2413.5; DB 1; Length 459;
Matches 43%; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MIVFVRFNSSYGFPEVEVDSDTSLIQLEKVAKRGVPAQDLRVIFAGKELPNHLTVQNC 60
DB 1 MIVFVRFNSSYGFPEVEVDSDTSLIQLEKVAKRGVPAQDLRVIFAGKELPNHLTVQNC 60
QY 61 LEQOSIVHIVQRPQRSHETHNASGGDEPOSTSEGSIMESRLTPVLSLSSHTLPVDSVGLA 120
DB 61 LEQOSIVHIVQRPQRSHETHNASGGDEPOSTSEGSIMESRLTPVLSLSSHTLPVDSVGLA 120
QY 121 VILDTDSKSDSAAARGP-VKPTNSFFTYCKGPKHKVQPKLRVQCGTCQKATLTLAGGP 179
DB 121 VILDTDSKSDSAAARGP-VKPTNSFFTYCKGPKHKVQPKLRVQCGTCQKATLTLAGGP 179
QY 180 SCMDVLIIPNRMSGSCSPDCGPTGTRAFEFKCGAPPSDKOTSVALNLTNSRSIPCIA 239
DB 180 SCMDVLIIPNRMSGSCSPDCGPTGTRAFEFKCGAPPSDKOTSVALNLTNSRSIPCIA 239
QY 240 CTDVSPVLVPCNRRHVICLDCFHLVCTRLNDRQFVDAQLGYSLPCVAGCNSLIKE 300
DB 240 CTDVSPVLVPCNRRHVICLDCFHLVCTRLNDRQFVDAQLGYSLPCVAGCNSLIKE 300
QY 300 LHFRLILBEQVTRYQOYGABECVLMGSGVLCPRPGCAGLLPECGQKVKTCGGNGLGC 359
DB 300 LHFRLILBEQVTRYQOYGABECVLMGSGVLCPRPGCAGLLPECGQKVKTCGGNGLGC 359
QY 360 GFVFCRCCKEAYHEGDCSLLEPSGATSCAYVQKRAAEQARWEASKEITIKTKPCPR 419
DB 360 GFVFCRCCKEAYHEGDCSLLEPSGATSCAYVQKRAAEQARWEASKEITIKTKPCPR 419
QY 420 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 458
DB 420 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 458
QY 421 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 459
DB 421 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 459

RESULT 6
QK5C3 PRELIMINARY; PRT; 437 AA.
AC QK5C3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parkin isoform.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RJ Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF381278; AAM2455.1;
DR InterPro; IPR003977; parkin.
DR InterPro; IPR006266; Ubiquitin.
DR InterPro; IPR002867; Znf_C6HC.
DR Pfam; PF01485; IER; 1.
DR PRINTS; PRO1475; PARKIN.
DR PRINTS; PRO0348; UBICUTIN.
DR SMART; SMO0647; IER; 2.
DR SMART; SMO0213; UBQ; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00553; UBICUTIN_2; 1.
SQ SEQUENCE 446 AA; 49368 MW; F06FAF788531DF5C CRC64;

Query Match
Best Local Similarity 86.7%; Score 2203; DB 1; Length 446;
Matches 41%; Conservative 12; Mismatches 12; Indels 29; Gaps 2;

DR SMART; SMO0213; UBQ; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00553; UBICUTIN_2; 1.
SQ SEQUENCE 437 AA; 48734 MW; 3227286E41B5E5E3 CRC64;

Query Match
Best Local Similarity 88.6%; Score 2265.5; DB 1; Length 437;
Matches 41%; Conservative 12; Mismatches 12; Indels 29; Gaps 2;

QY 1 MIVFVRFNSSYGFPEVEVDSDTSLIQLEKVAKRGVPAQDLRVIFAGKELPNHLTVQNC 60
DB 1 MIVFVRFNSSYGFPEVEVDSDTSLIQLEKVAKRGVPAQDLRVIFAGKELPNHLTVQNC 60
QY 61 LEQOSIVHIVQRPQRSHETHNASGGDEPOSTSEGSIMESRLTPVLSLSSHTLPVDSVGLA 120
DB 61 LEQOSIVHIVQRPQRSHETHNASGGDEPOSTSEGSIMESRLTPVLSLSSHTLPVDSVGLA 120
QY 121 VILDTDSKSDSAAARGP-VKPTNSFFTYCKGPKHKVQPKLRVQCGTCQKATLTLAGGP 179
DB 121 VILDTDSKSDSAAARGP-VKPTNSFFTYCKGPKHKVQPKLRVQCGTCQKATLTLAGGP 179
QY 180 SCMDVLIIPNRMSGSCSPDCGPTGTRAFEFKCGAPPSDKOTSVALNLTNSRSIPCIA 239
DB 180 SCMDVLIIPNRMSGSCSPDCGPTGTRAFEFKCGAPPSDKOTSVALNLTNSRSIPCIA 239
QY 240 CTDVSPVLVPCNRRHVICLDCFHLVCTRLNDRQFVDAQLGYSLPCVAGCNSLIKE 300
DB 240 CTDVSPVLVPCNRRHVICLDCFHLVCTRLNDRQFVDAQLGYSLPCVAGCNSLIKE 300
QY 300 LHFRLILBEQVTRYQOYGABECVLMGSGVLCPRPGCAGLLPECGQKVKTCGGNGLGC 359
DB 300 LHFRLILBEQVTRYQOYGABECVLMGSGVLCPRPGCAGLLPECGQKVKTCGGNGLGC 359
QY 360 GFVFCRCCKEAYHEGDCSLLEPSGATSCAYVQKRAAEQARWEASKEITIKTKPCPR 419
DB 360 GFVFCRCCKEAYHEGDCSLLEPSGATSCAYVQKRAAEQARWEASKEITIKTKPCPR 419
QY 420 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 458
DB 420 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 458
QY 421 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 459
DB 421 CNVPIEKNGGGMHKCPQPOCKLEMCNCGCENRACMG 459

RESULT 7
QK5C3 PRELIMINARY; PRT; 446 AA.
AC QK5C3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parkin isoform.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RJ Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF381278; AAM2455.1;
DR InterPro; IPR003977; parkin.
DR InterPro; IPR006266; Ubiquitin.
DR InterPro; IPR002867; Znf_C6HC.
DR Pfam; PF01485; IER; 1.
DR PRINTS; PRO1475; PARKIN.
DR PRINTS; PRO0348; UBICUTIN.
DR SMART; SMO0647; IER; 2.
DR SMART; SMO0213; UBQ; 1.
SQ SEQUENCE 446 AA; 49368 MW; F06FAF788531DF5C CRC64;

Query Match
Best Local Similarity 86.7%; Score 2203; DB 1; Length 446;
Matches 41%; Conservative 12; Mismatches 12; Indels 29; Gaps 2;

```

Best Local Similarity 93.4%, Score 210.5, DB 4, Length 465;
Matches 405; Conservative 14; Mismatches 14; Indels 2; Gaps 2;

QY 1 MIVFVRNNSGYPFVEVDSOTSILOLKEVYAKQGVADQRLVIFAGKELPHILTVQNC 60
DB 1 MIVFVRNNSGYPFVEVDSOTSILOLKEVYAKQGVADQRLVIFAGKELPHILTVQNC 60
QY 61 LEQOSIVHIVORPRRSHETNASGGDEPOSTSEGSIMESRESTRVLSHSLPVSQVLA 120
DB 61 LEQOSIVHIVORPRRSHETNASGGDEPOSTSEGSIMESRESTRVLSHSLPVSQVLA 120
QY 121 VILDTCKRDEAARGV-KPTVNSFFIKYKCHVQPKLAVQCGTCKCAT-TLAQGP 179
DB 121 VILDTCKRDEAARGV-KPTVNSFFIKYKCHVQPKLAVQCGTCKCAT-TLAQGP 179
QY 180 SCMDVILPNRMSGECQSPCPGTRAEFFKCGAHTSDKDTSVALKLITNSRSIPCA 239
DB 180 SCMDVILPNRMSGECQSPCPGTRAEFFKCGAHTSDKDTSVALKLITNSRSIPCA 240
QY 240 CTQVRSPVLVFCQNRHVICLDCPHLYCVTRLNDRQFVHDAQGLYSLPCVAGCNSLIKE 299
DB 240 CTQVRSPVLVFCQNRHVICLDCPHLYCVTRLNDRQFVHDAQGLYSLPCVAGCNSLIKE 300
QY 300 LHPRIILGEEQYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGRKVTCEGNGILGC 359
DB 300 LHPRIILGEEQYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGRKVTCEGNGILGC 360
QY 360 GFVFCRCCKEAYHEGECDSLEPSGATSOAYRVYDKAAEQARWEASKETIKTKTKPCR 419
DB 360 GFVFCRCCKEAYHEGECDSLEPSGATSOAYRVYDKAAEQARWEASKETIKTKTKPCR 420
QY 420 CNVPIKNGGCVHMKCPCPOCKLEWKNKCCGKXACMGQHFV 464
DB 420 CNVPIKNGGCVHMKCPCPOCKLEWKNKCCGKXACMGQHFV 465

RESULT 8
Q60260 PRELIMINARY; PRT; 465 AA.
ID AC O60260
DT 01-AUG-1998 (TREMBLrel: 07, Created;
DT 01-AUG-1998 (TREMBLrel: 07, Last sequence update)
DT 01-MAR-2003 (TREMBLrel: 23, Last annotation update)
DE Parkin.
CN PARKIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
FN [.]
RP SEQUENCE FROM N.A.
RX MEDLINE=98219084; PubMed=9560156;
RA Kitada T., Asakawa S., Hattori N., Matsumine H., Yamamura Y.,
RA Miroshima S., Yokochi M., Mizuno Y., Shimizu Y.;
RT "Mutations in the parkin gene causes autosomal recessive juvenile
RT parkinsonism";
RL Nature 392:605-608(1998).
RL EMBL: AB009973; BAA25751.1; .
DR HSSP: P22248; IUB.
DR Genew: HGNC:8607; PARK2.
DR InterPro: IPR003977; Parkin.
DR InterPro: IPR000169; SHPOT acsite.
DR InterPro: IPR000626; Ubiquitin.
DR InterPro: IPR002867; Znf_C6HC.
DR Pfam: PF01485; IPR; 1.
DR Pfam: PF00240; Ubiquitin; 1.
DR PRINTS: PS01475; PARKIN.
DR PRINTS: PR00348; UBIQUITIN.
DR SMART: SM00647; IPR; 2.
DR SMART: SK0213; UBC; 1.
DR PROSITE: PS00139; THIOLE PROTEASE CY5; 1.
DR PROSITE: PS00053; UBIQUITIN.2; .
SQ SEQUENCE 465 AA; 51650 MW; 6A8B8989B76ECCAE CRC64;

Query Match 94.5%, Score 210.5, DB 4, Length 465;
Best Local Similarity 83.4%, Pred. No. 5, e-136;
Matches 369; Conservative 29; Mismatches 47; Indels 2; Gaps 1;

QY 1 MIVFVRNNSGYPFVEVDSOTSILOLKEVYAKQGVADQRLVIFAGKELPHILTVQNC 60
DB 1 MIVFVRNNSGYPFVEVDSOTSILOLKEVYAKQGVADQRLVIFAGKELPHILTVQNC 60
QY 61 LEQOSIVHIVORPRRSHETNASGGDEPOSTSEGSIMESRESTRVLSHSLPVSQVLA 120
DB 61 LEQOSIVHIVORPRRSHETNASGGDEPOSTSEGSIMESRESTRVLSHSLPVSQVLA 120
QY 121 VILDTCKRDEAARGV-KPTVNSFFIKYKCHVQPKLAVQCGTCKCAT-TLAQGP 179
DB 121 VILDTCKRDEAARGV-KPTVNSFFIKYKCHVQPKLAVQCGTCKCAT-TLAQGP 179
QY 180 SCMDVILPNRMSGECQSPCPGTRAEFFKCGAHTSDKDTSVALKLITNSRSIPCA 239
DB 180 SCMDVILPNRMSGECQSPCPGTRAEFFKCGAHTSDKDTSVALKLITNSRSIPCA 240
QY 240 CTQVRSPVLVFCQNRHVICLDCPHLYCVTRLNDRQFVHDAQGLYSLPCVAGCNSLIKE 299
DB 240 CTQVRSPVLVFCQNRHVICLDCPHLYCVTRLNDRQFVHDAQGLYSLPCVAGCNSLIKE 300
QY 300 LHPRIILGEEQYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGRKVTCEGNGILGC 359
DB 300 LHPRIILGEEQYTRYQYGAEECVLQMGVLCPRPGCGAGLLPEQGRKVTCEGNGILGC 360
QY 360 GFVFCRCCKEAYHEGECDSLEPSGATSOAYRVYDKAAEQARWEASKETIKTKTKPCR 419
DB 360 GFVFCRCCKEAYHEGECDSLEPSGATSOAYRVYDKAAEQARWEASKETIKTKTKPCR 420
QY 420 CNVPIKNGGCVHMKCPCPOCKLEWKNKCCGKXACMGQHFV 464
DB 420 CNVPIKNGGCVHMKCPCPOCKLEWKNKCCGKXACMGQHFV 465

RESULT 9
Q6W007 PRELIMINARY; PRT; 357 AA.
ID AC Q6W007
DT 01-MAR-2002 (TREMBLrel: 23, Created)
DT 01-MAR-2002 (TREMBLrel: 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel: 23, Last annotation update)
DE Similar to Parkinson disease (Autosomal recessive, juvenile); 2.
DE parkin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
FN [.]
RP SEQUENCE FROM N.A.
RX TISSUES=Testis;
FA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC022614; AAB22014.1; .
DR InterPro: IPR003977; Parkin.
DR InterPro: IPR005626; Ubiquitin.
DR InterPro: IPR002867; Znf_C6HC.
DR Pfam: PF01485; IPR; 1.
DR Pfam: PF00240; Ubiquitin; 1.
DR Pfam: PF04810; Zf-Sec23_Sec24; 1.
DR PRINTS: PS01475; PARKIN.
DR PRINTS: PR00348; UBIQUITIN.
DR SMART: SM00213; UBC; 1.
DR PROSITE: PS00053; UBIQUITIN.2; 1.
SQ SEQUENCE 357 AA; 42407 MW; 48C0F41C8622A5C6 CRC64;

Query Match 62.4%, Score 1604; DB 4; Length 367;
Best Local Similarity 73.0%, Pred. No. 1e-142;
Matches 301; Conservative 21; Mismatches 39; Indels 20; Gaps 2;

```

QY 1 MIVFVFNSSYGFPPVEVDSDTSLQKKEVAKQGVADQLRVIFAGKELPNHLITVQCD 60
DB 1 MIVFVFNSSHGFPVEVDSDTSLQKKEVAKQGVADQLRVIFAGKELPNHLITVQCD 60
QY 61 LEOOSTVHV:VQSPRSHSTNAGGDEPCSTSEGSIKESRS:TRVD:SSHTLP:VDVSLA 120
DB 61 LEOOSTVHV:VQSPRSHSTNAGGDEPCSTSEGSIKESRS:TRVD:SSHTLP:VDVSLA 120
QY 121 VILDTDSKRDSEAAKGPV:KPTYNFFIYKGPCHKVGPKLRYQCGCKQATLITLAQGP 179
DB 121 VILDTDSKRDSEAAKGPV:KPTYNFFIYKGPCHKVGPKLRYQCGCKQATLITLAQGP 179
QY 121 VILHTDSRSDSPAGSPACRSIYNFVYKGPQGVQPKLVQVQSTCRCATLITLQGP 120
DB 121 VILHTDSRSDSPAGSPACRSIYNFVYKGPQGVQPKLVQVQSTCRCATLITLQGP 120
QY 180 SCHDDVLI:INRMSGSCQSPDCSTGTRAEFFKCGAHTPSDKTISVALNLTISNRSIPCTA 239
DB 180 SCHDDVLI:INRMSGSCQSPDCSTGTRAEFFKCGAHTPSDKTISVALNLTISNRSIPCTA 239
QY 181 SCHDDVLI:INRMSGSCQSPDCSTGTRAEFFKCGAHTPSDKTISVALNLTISNRSIPCTA 240
DB 181 SCHDDVLI:INRMSGSCQSPDCSTGTRAEFFKCGAHTPSDKTISVALNLTISNRSIPCTA 240
QY 240 CTIDVRSPLVFCQNRHRRVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP: 289
DB 240 CTIDVRSPLVFCQNRHRRVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP: 289
QY 241 CTIDVRSPLVFCQNRHRRVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP: 300
DB 241 CTIDVRSPLVFCQNRHRRVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP: 300
QY 289 -----VAGCPNSL:KELHFR:ILGBEQYTRYQYGAECVLMQGVLP:PRGCGAGL 340
DB 289 -----VAGCPNSL:KELHFR:ILGBEQYTRYQYGAECVLMQGVLP:PRGCGAGL 340
QY 301 ALGGFRGAGGCPNSL:KELHFR:ILGBEQYTRYQYGAECVLMQGVLP:PRGCGAGL 360
DB 301 ALGGFRGAGGCPNSL:KELHFR:ILGBEQYTRYQYGAECVLMQGVLP:PRGCGAGL 360
QY 341 LPEQGRKVT:CEGNSLGGCGF 361
DB 341 LPEQGRKVT:CEGNSLGGCGF 361
QY 361 LPEQGRKVT:CEGNSLGGCGF 381
DB 361 LPEQGRKVT:CEGNSLGGCGF 381

```

RESULT 10

```

ID Q8VHV6 PRELIMINARY: PRT: 274 AA.
AC Q8VHV6
DT 01-MAR-2002 (TRENDELrel. 20, Created:
DT 01-MAR-2002 (TRENDELrel. 20, Last sequence update:
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update:
DE Parkin transcript variant 6 (Parkin isoform).
CS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1;
RS SEQUENCE FROM N.A.
RA D'Agata V.M., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin in the rat brain.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RP 1;
RC STRAIN=Sprague-Dawley;
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF343574; AAL73348.1; -
DR EMBL; AF381280; AAM21455.1; -
DR InterPro; IPR003977; parkin.
DR InterPro; IPR003977; parkin.
DR Pfam; PFO1485; 1BR; 1.
DR PRINTS; PR01475; PARKIN.
DR PROSITE; PS00119; THIOLE PROTEASE CYS. 1.
SQ SEQUENCE 274 AA; 30642 MW; B29BB58E8E5BDFE CRC64;

```

```

Query Match 59.6%; Score 1512; DB 1; Length 274;
Best Local Similarity 96.4%; Pred. No. 4.2e-136;
Matches 264; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 19: YSGCQSPDCPTRAEPRFFKCGAHTPSDKTISVALNLTISNRSIPCTIACDVRSPLVVF 250
DB 1: MSCEQSPDCPTRAEPRFFKCGAHTPSDKTISVALNLTISNRSIPCTIACDVRSPLVVF 60
QY 251 QCNHRHVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP:CVAGCPNSL:KELHFR:ILGEEQ 310
DB 251 QCNHRHVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP:CVAGCPNSL:KELHFR:ILGEEQ 310

```

```

DB 61 QCNHRHVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP:CVAGCPNSL:KELHFR:ILGEEQ 120
QY 31: YTRYQYGAECVLMQGVLP:PRGCGAGL:PEQGRKVT:CEGNSLGGCGF:VFCRCKEA 370
DB 12: YTRYQYGAECVLMQGVLP:PRGCGAGL:PEQGRKVT:CEGNSLGGCGF:VFCRCKEA 180
QY 371 YHEGDCDLSLLEPSGATSCAYRVDRKAAEQARWEASAKETIKTKTPCRCNVPIFKAGGC 430
DB 181 YHEGDCDLSLLEPSGATSCAYRVDRKAAEQARWEASAKETIKTKTPCRCNVPIFKAGGC 430
QY 431 YHMKCPQPCCKLEKWCNCCGCEWNRACNCHDHWFV 464
DB 241 YHMKCPQPCCKLEKWCNCCGCEWNRACNCHDHWFV 274

```

RESULT 11

```

ID Q8NI44 PRELIMINARY: PRT: 274 AA.
AC Q8NI44
DT 01-OCT-2002 (TRENDELrel. 22, Created:
DT 01-OCT-2002 (TRENDELrel. 22, Last sequence update:
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update:
DE Parkin isoform.
CS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A.
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF381282; AAM21455.1; -
DR InterPro; IPR003977; parkin.
DR InterPro; IPR003977; parkin.
DR Pfam; PFO1485; 1BR; 1.
DR PRINTS; PR01475; PARKIN.
DR SMART; SMC0647; 1BR; 2.
DR PROSITE; PS00119; THIOLE PROTEASE CYS. 1.
SQ SEQUENCE 274 AA; 30626 MW; 3A5SD26C14A8:7CC CRC64;

```

```

Query Match 55.3%; Score 1420; DB 4; Length 274;
Best Local Similarity 88.3%; Pred. No. 1.5e-125;
Matches 242; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 19: MSCEQSPDCPTRAEPRFFKCGAHTPSDKTISVALNLTISNRSIPCTIACDVRSPLVVF 250
DB 1: MSCEQSPDCPTRAEPRFFKCGAHTPSDKTISVALNLTISNRSIPCTIACDVRSPLVVF 60
QY 251 QCNHRHVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP:CVAGCPNSL:KELHFR:ILGEEQ 310
DB 61 QCNHRHVICLDFPHLYCVTR:LNDRQFVHDAQ:GYS:LP:CVAGCPNSL:KELHFR:ILGEEQ 120
QY 31: YTRYQYGAECVLMQGVLP:PRGCGAGL:PEQGRKVT:CEGNSLGGCGF:VFCRCKEA 370
DB 12: YTRYQYGAECVLMQGVLP:PRGCGAGL:PEQGRKVT:CEGNSLGGCGF:VFCRCKEA 180
QY 371 YHEGDCDLSLLEPSGATSCAYRVDRKAAEQARWEASAKETIKTKTPCRCNVPIFKAGGC 430
DB 181 YHEGDCDLSLLEPSGATSCAYRVDRKAAEQARWEASAKETIKTKTPCRCNVPIFKAGGC 430
QY 431 YHMKCPQPCCKLEKWCNCCGCEWNRACNCHDHWFV 464
DB 241 YHMKCPQPCCKLEKWCNCCGCEWNRACNCHDHWFV 274

```

RESULT 12

```

ID Q8ES22 PRELIMINARY: PRT: 255 AA.
AC Q8ES22
DT 01-MAR-2001 (TRENDELrel. 16, Created:
DT 01-MAR-2001 (TRENDELrel. 16, Last sequence update:

```

```

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Parkin.
GN PARK2.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BALB/c; Tissue=Kidney;
RA Stichel C.C., Augustin M., Kuhn K., Zhu X., Engels P., Ulmer C.,
RA Lubbert H.;
RL "Parkin expression in the adult mouse brain.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF250295; AAG13892.1; -
DR HSP; P02248; IJBI.
DR WGD; MGI:1355296; Park2.
DR InterPro; IPRO03977; Parkin.
DR Pfam; PF02240; Ubiquitin; 1.
DR PRINTS; PR01475; PARKIN.
DR SMART; SW00213; UBC; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 255 AA; 28155 MW; D2EEC07EC9C4C1A CRC64;

Query Match 49.9%; Score 1281.5; DB 11; Length 255;
Best Local Similarity 99.6%; Pred. No. 1.6e-112;
Matches 244; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MIVFVFNSSYGPPVEVDSOTS:LQLKEVAKROGVADQLRVIFAGKELPNHLYVQNC 60
DB 1 MIVFVFNSSYGPPVEVDSOTS:LQLKEVAKROGVADQLRVIFAGKELPNHLYVQNC 60

QY 61 LEQOSIVHIVORPRRSHEFTNASGGDEPOSTSGSINERSLTRVJSSHTLVDSVGLA 120
DB 61 LEQOSIVHIVORPRRSHEFTNASGGDEPOSTSGSINERSLTRVJSSHTLVDSVGLA 120

QY 121 VILDTSKRSEAAAGP-VKPTYNFFIYCKGPKHVQKLRVQCGTCKQATLTLAQGP 179
DB 121 VILDTSKRSEAAAGP-VKPTYNFFIYCKGPKHVQKLRVQCGTCKQATLTLAQGP 179

QY 180 SCWDDVLINRMSEGCSPDCPTRAEFKCGAHTPSDKDTSVALNLTNSRRSIPICIA 219
DB 181 SCWDDVLINRMSEGCSPDCPTRAEFKCGAHTPSDKDTSVALNLTNSRRSIPICIA 243

QY 240 CTQVR 244
DB 241 CTQVR 245

RESULT 13
ID Q95E23 PRELIMINARY; PRT; 262 AA.
AC Q95E23;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Parkin.
GN PARK2.
OS M-s musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BALB/c; Tissue=Brain;
RA Stichel C.C., Augustin M., Kuhn K., Zhu X., Engels P., Ulmer C.,
RA Lubbert H.;
RL "Parkin expression in the adult mouse brain.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF250294; AAG13891.1; -
DR HSP; P02248; IJBI.

```

```

DR WGD; MGI:1355296; Park2.
DR InterPro; IPRO03977; Parkin.
DR Pfam; PF02240; Ubiquitin; 1.
DR PRINTS; PR01475; PARKIN.
DR SMART; SW00213; UBC; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 262 AA; 28700 MW; CDC86CP2956551AE CRC64;

Query Match 49.7%; Score 1276.5; DB 11; Length 262;
Best Local Similarity 99.6%; Pred. No. 4.8e-112;
Matches 243; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MIVFVFNSSYGPPVEVDSOTS:LQLKEVAKROGVADQLRVIFAGKELPNHLYVQNC 60
DB 1 MIVFVFNSSYGPPVEVDSOTS:LQLKEVAKROGVADQLRVIFAGKELPNHLYVQNC 60

QY 61 LEQOSIVHIVORPRRSHEFTNASGGDEPOSTSGSINERSLTRVJSSHTLVDSVGLA 120
DB 61 LEQOSIVHIVORPRRSHEFTNASGGDEPOSTSGSINERSLTRVJSSHTLVDSVGLA 120

QY 121 VILDTSKRSEAAAGP-VKPTYNFFIYCKGPKHVQKLRVQCGTCKQATLTLAQGP 179
DB 121 VILDTSKRSEAAAGP-VKPTYNFFIYCKGPKHVQKLRVQCGTCKQATLTLAQGP 180

QY 180 SCWDDVLINRMSEGCSPDCPTRAEFKCGAHTPSDKDTSVALNLTNSRRSIPICIA 239
DB 181 SCWDDVLINRMSEGCSPDCPTRAEFKCGAHTPSDKDTSVALNLTNSRRSIPICIA 240

QY 240 CTQV 243
DB 241 CTQV 244

RESULT 14
ID Q95T14 PRELIMINARY; PRT; 482 AA.
AC Q95T14;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Drosophila melanogaster (Fruit fly).
GN PARKIN; CR CG:50523.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscoidiptera;
OX Ephydroidea; Drosophilidae; Drosophila.
RN [1]_TaxID=7227;
RP SEQUENCE FROM N.A.
RA Stadelman M., Brokstein P., Hong J., Asabayani A., Carlson J.,
RA Campe X., Chavez C., Corsett V., Farfan S., Frise S., George R.,
RA Gonzalez M., Guarin H., Li P., Diao G., Miranda A., Wangli C.,
RA Nuno J., Paolel J., Paragas V., Park S., Phouatavong S., Wan K.,
RA Xu C., Lewis S.E., Rubin G.M., Celis Kex S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY058754; AAL13983.1; -
DR FlyBase; FBgn034150; parkin.
DR InterPro; IPRO03577; N6 Mtase.
DR InterPro; IPRO03577; parkin.
DR InterPro; IPRO03626; Ubiquitin.
DR Pfam; PF01485; UBR; 1.
DR Pfam; PF0245; Ubiquitin; 1.
DR PRINTS; PR01475; PARKIN.
DR PRINTS; PR03459; UBIQUITIN.
DR SMART; SMC0647; UBR; 2.
DR SMART; SM00213; UBC; 1.
DR PROSITE; PS03092; N6 MTASE; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 482 AA; 54139 MW; F0A154.F6D748D07 CRC64;

```

```

Query Match          41.9%; Score 1078; DB 5; Length: 482;
Best Local Similarity 43.7%; Pred. No. 5,3e-93;
Matches 209; Conservative 78; Mismatches 145; Indels 46; Gaps 12;

QV 3 VVFVNSYGFPPVDSVDSSTSLQKKEVYAKRGVAPDQALRVFAGKELPHNLTVNCDELE 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2b 32 IYVXTNKTITVLEQWDIKNVKEJVAPQLGSPDDKLIIFAGKELSDATTEECCLG 91
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 QQS-VH-VQ-RP-RRSHSHETNAGSGDEPOSTBEGSTWESRS-TRYDLASHLTPVDSGL 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 QQSVLHAIRLAPPVQROKIQSATLEEBEPSLSDEAS-----KPLNETLL 135
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 AVILDTSK-----RSEARGPVKTYNSFYCKGPKHKVQPKLKVQCTKCATJILA 176
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 DLQLESEELNITDEERYAKA-----HFFVHC-SQCDKLCNKGKLVRCALCKGGAFYH 189
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 QGPGCWDVLLPNRYSGEQCPD--CFGR-----AEFFPKGAFHT-SDKDTSVALKL 227
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 RDECKWDVLSRR:PGHCESELVACVDNAGDPFPAEFERFKAENVSGCKDFAPJNL 249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 ITSRRSIPQACTDVRSPVLVPCNHRHVICLDCPHLYCVTELRNDFQVHDAGLOYSLP 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 IKNNFKVPCVCLACTDVSQTVVFFCASQHTVC:DCEFRHYCRSLRGRQFMPPDFCYTLP 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 CVASCPNSLIXELHFRILGEEQVTRYCOYGAEBCV:QMSGVLCPRPGGAGLLPFCQGR 347
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 CPACGSHSFIEE:HFHKLITREEDVDRQRTATEBYVLQAGVLCVPGCGXGLLVBPCCR 369
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 KVTGEGNGLCGVFCRDCKEAYHEGDCUSLLEFSQAT---SQAYVDPKRAAEQARNEE 404
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 KVTQCN-----CGGVVFCRNC:QGVHIGEC--LPBGTCASATNSCEVTVDPNRAAEARWD 423
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 ASKETIKKTTPCRCNVPIEKNGGCHMKCPQCKLEWNCWCGEWRACGHDWF 462
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 ASNVTIKVTSPCPKCRKTFETROCGGEMVCTRAGCGFECWCVCQETWRDCNGAHWF 491
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q9VP72
ID C9VP72      EMBL:INRA; PRT: 492 AA.
AC Q9VP72.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
ST 01-MAR-2003 (TrEMBLrel. 22, Last sequence update)
DE CG10523 protein (PARKIN).
GN PARKIN OR CG10523.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beckley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beron P.V., Bertram B.P., Brandari D., Bolshakov S.,
RA Borokova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busar D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dalke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.K.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,

```

```

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison C.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D., Nelson D.B.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paczes C.M.,
RA Palazzolo M., Pittman K.S., Pan S., Pollard J., Puti V., Reese V.G.,
RA Reinert K., Remington K., Saunders P.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Stadling A.C., Stapleton M., Strong A., Sun E.,
RA Svraskas R., Tector C., Turner R., Venter G., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhao Y., Zhou G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "the genome sequence of Drosophila melanogaster";
RA Science 287:2195-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC Celisner S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Batzer J., An H., Baldwin D., Baizen J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.K.,
RA Dodson K., Doup L.E., Doyle C., Dresnek P., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzales M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jallali M., Kuse D., Li F., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nelson K.A., Nimmo J.,
RA Paules J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Proudenavong S., Pittman G.S., Puti V., Richards S., Scheeler F.,
RA Stapleton M., Strong A., Svraskas R., Tector C., Tyler C.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
RA "Sequencing of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Xisara S., Crosby M.A., Matthews S.E., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker S.S., Prochuk S.E., Smith C.O.,
RA Tupy J., Bergman B., Bertram B., Carlson J.W., Celisner S.E.,
RA Clapp X., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Seale S.M., Smith E., Shu S., Smutnick F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
RA "Annotation of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter J.C.,
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Kang S.C., Ahn J., Bae Y., Lee Z., Park K.,
RA "Characterization of Parkin in Drosophila melanogaster";
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Bae Y., Kang S., Park K.,
RA "Genome Structure Analysis and Characterization of Parkin in
RA Drosophila melanogaster";
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP EMBL; A030593; AAF51831.2;
DR EMBL; A030593; AAF51831.2;
DR EMBL; AY093423; AAM18800.2;
DR EMBL; AF510072; AAM43910.1;
DR EMBL; P02248; IUPI.
DR FlyBase; FBgn041120; Parkin.
DR InterPro; IPR020552; N6_Ytase.

```

```
DR InterPro; IPR003977; parkin.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR002867; Znf_C6HC.
DR Pfam; PF01485; IER; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR01475; PARKIN.
DR PRINTS; PR00348; UBQUITIN.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
DR PROSITE; PS00053; UBQUITIN 2; 1.
SQ SEQUENCE 482 AA; 54104 MW; 4A89C695475B2113 CRC64;

Query Match 41.9%; Score 1079; DB 5; Length 482;
Best Local Similarity 43.7%; Pred. No. 5.3e-93;
Matches 209; Conservative 79; Mismatches 145; Indels 46; Gaps 12;

QY 3 VEVRENSYGFPEVDESDTSILQKEVVAKSGVGPADQLRVIFAGKELPNHLTVQCDEL 62
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 32 IVVKTNTGKTLVNLPCWDINVKELVAPQLGLOPDLXIIIFAGKELSDATTIEQCDLG 91
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 QGSIVHIVQ-RP--RRRSHE--NASGGDEPQSTSGSIWESRSLTRVLSSTHTUPVLSVL 119
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 QGSVLHARLRPPVGRQKIQATLEEEEPSLSDEAS-----KPLNETLL 135
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 AVILDTQK---RQSEAAARGPVKPTYNSEFFIYKGPCHKVPQKLRVCCGCTCKQATLLA 176
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 DLQLESEERLNTDEERVAKA-----HFFVHC--SQCDKLCNGKLRVRCALCKGGAFTVH 189
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 QGPSQNDVLIPIRMSGCGQSPD--CPGTR-----AEFFKCGAHT-SQKTSVALNL 227
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 RDEPCWDVLRRIIPGHCESEVACVNAAGDPFFAEFFKCAEHVSGGKGFAPDLNL 249
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 ITSNRSIPCIACDVRSPVAVFCNHRHVICLCPHLYCVYVTRLNDRQFVYDAGLGYSILP 287
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 IKNNIKWVCLACTDVSDTLVFPCCASCHV-CIDCFHYCRLGERCFMHPDFCYTLF 309
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 288 CVAGCNSLILKELHFRILGSEQYTRYQYQYGAECVLQMSGVLCPRGCGAGLLPECGOR 347
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 CPAGCEHSFTEIHFHFKLLTREVDRYGRFATERYVJAGGV-CPQGGQMGLLVSPCCR 369
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 KVTCEGNGLGCGFYFCRDCKEAVHEGDCDGLLEPSGAT---SOAYRVYDRAAEQARWEE 404
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 KVTCCN---GQGVYFCNCLQGYHIGC--LPSTGASATNSCEYTVDFNRRAEAKWE 423
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 405 ASKETKKTKPCPCNVPIENKGCYHMKCPQPCCKLEKWCNCGGCWNRACVGDHWF 462
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 424 ASNVTKVSKPCPKCTPTEDGGCMHMCVTRAGCGGFEMCWVQCQTEWTRDCMGHWF 481
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: October 21, 2003, 21:51:05
Job time: 154 secs